Renormalization of the Coulomb Interaction in the Hubbard Model and in Dynamical Mean Field Theory

von
Clemens Adolphs

Diplomarbeit in PHYSIK

vorgelegt der
Fakultät für Mathematik, Informatik und Naturwissenschaften
der Rheinisch-Westfälischen Technischen Hochschule Aachen

im
Oktober 2010

angefertigt an der
German Research School for Simulation Sciences

bei
Prof. Dr. E. Koch

Contents

1. Introduction

While in the realm of particle physics and general relativity the precise nature of the very laws of physics is all but understood, condensed matter physicists are in possession of a theory of (almost) everything. As pointed out by Dirac [1], all we need to know of a system is its wavefunction $|\Psi\rangle$. This is obtained by solving the Schrödinger equation

$$H\,|\Psi\rangle = E\,|\Psi\rangle \tag{1.1}$$

for the Hamiltonian of electrons and nuclei,

$$H = -\sum_{\alpha=1}^{N_n} \frac{\mathbf{P}_\alpha^2}{2M_\alpha} - \sum_{j=1}^{N_e} \frac{\mathbf{p}_j^2}{2m} - \sum_{j=1}^{N_e}\sum_{\alpha=1}^{N_n} \frac{Z_\alpha e^2}{|\mathbf{r}_i - \mathbf{R}_\alpha|} + \sum_{j<k}^{N_e} \frac{e^2}{|\mathbf{r}_j - \mathbf{r}_k|} + \sum_{\alpha<\beta}^{N_n} \frac{Z_\alpha Z_\beta e^2}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|}, \tag{1.2}$$

where N_e, $\mathbf{p}_j$ and $\mathbf{r}_j$ denote number, momentum and position of the electrons, while N_n, $\mathbf{P}_\alpha$ and $\mathbf{R}_\alpha$ are the same for the nuclei, whose atomic numbers are the Z_α. From a solid-state physicist's point of view, these are the fundamental building blocks of the material world. One might object that such a Hamiltonian cannot possibly be a theory of everything, because it completely ignores all nuclear and sub-nuclear degrees of freedom. Should not a theory of everything start from quantum chromodynamics (QCD)? In principle, yes. However, excitations in these realms are energetically much higher than all electronic excitations in solids. Consequently, their dynamics are much faster and can safely be integrated out. Moreover, perturbation theory tells us that corrections to energies and wavefunctions scale with $\mathcal{O}(1/\Delta E)$ where ΔE is the difference in energies. Fig. 1.1 illustrates some energy scales in physics. All the energetically higher excitations do is to renormalize the bare parameters of our *effective* model. For example, the anomalous magnetic moment of the electron is a consequence of quantum electrodynamics (QED) and its mass is hypothesized to be a consequence of the Higgs mechanism. The precise nature of the underlying mechanisms leading to mass, charge, and spin of the electron, however, are largely irrelevant for the dynamics of condensed matter systems. It only matters that these parameters are there and give rise to an (effective) Hamiltonian (1.2).

The wavefunction contains all information about a quantum state. Hence, from a solution of the Schrödinger equation one should, in principle, be able to derive all properties of a solid. The solution should predict optical and electrical properties as well as transition temperatures of phase transitions. Promising as this may sound, it is an endeavor doomed to fail. The first obstacle is that the computational resources required to diagonalize the Hamiltonian are on a truly astronomical scale. Imagine we just want to store the electronic wavefunction of a single iron atom with its 26 electrons. Even if we use a very crude grid of $10 \times 10 \times 10$ points in space only, we need an array of size $1000^{26} = 10^{78}$.

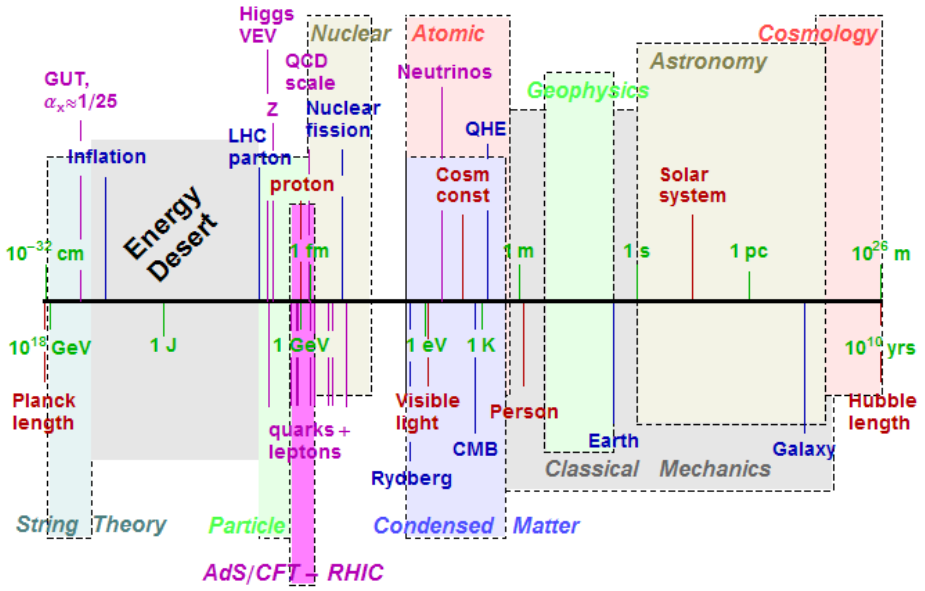


Figure 1.1.: Energy scales in physics. Condensed matter systems, dealing with electronic and phononic excitations, is well separated from nuclear and sub-atomic excitations. From http://www.princeton.edu/physics/research/high-energy-theory/gubser-group/introduction-to-the-physi/energy-scales-in-physics/

For adequate precision, we should use at least 16 bits. If we could store one bit using a single hydrogen atom, we would require 10^{79} atoms. This is on the same order of magnitude as current estimates for the number of atoms in the entire observable universe. And remember, this is for a single wavefunction of a single iron atom. For a whole chunk of iron, our memory requirements easily exceed 10^{100} bits.

The second objection is of a different kind. Let us suppose a relative of Laplace's daemon [2] could indeed calculate and store the entire wavefunction of a macroscopic system. Although he could accurately predict that glass is transparent for visible light and that copper is a good conductor, he could not come up with a generalizing explanation. In other words, his calculations will not tell him anything that simply doing the experiment won't. He has just repeated the experiment. What we mean when we talk about *understanding* a phenomenon is our ability to provide generalizations and give an explanation of the phenomenon that is on a higher level of abstraction than the phenomenon itself. All hints to these generalizing laws are drowned in the flood of information provided by the wavefunction. We see that the mere ability to reduce all the complex phenomena of solid state physics (and beyond) to the Hamiltonian (1.2) does not grant the ability to start from it and rebuild the world from scratch, as discussed in a famous essay by Anderson [3]. The bottom line of Anderson's essay *More is Different* is that among the hierarchy of ever more complex physical systems, entirely new properties and laws *emerge* that are not obvious from the underlying microscopic equations. In addition, it has been suggested that principles similar to those of Goedel's incompleteness theorem or Turing's halting problem put a limit on what can in principle be deduced from natural laws [4].

As was already pointed out by Dirac, the computational complexity of the many-body problem calls for approximations and models so that it can be solved efficiently. While approximations facilitate the calculation of results, the development of models is crucial to a thorough understanding of the relevant physics. Thus, models have to be chosen carefully, so that the reduced models still contain the physics one wants to study. In this chapter, we outline the problems one encounters in this process, and we introduce a particularly popular model that shall be the foundation of the studies in this thesis.

1.1. Ab-initio Methods and Model Hamiltonians

There are two fundamentally different ways to approach the many-body problem, corresponding to the two difficulties mentioned above. *Ab-initio* methods aim to directly simplify the problem of diagonalizing the Hamiltonian (1.2). This can be done, for example, by restricting the set of considered wavefunctions or by modifying the Hamiltonian. An example of the former is the Hartree-Fock method, where the wavefunction is assumed to be a single Slater determinant. An example of the latter is the free electron model, where the electron-electron interaction in (1.2) is replaced by an effective potential based on the single-electron picture. A very successful ab-initio method is Density Functional Theory [5, 6], which self-consistently maps the many-body Hamiltonian to an effective single-particle problem in the form of functionals of the electron density. While this method is theoretically exact, the true functional is unknown and thus has to be approximated. Nevertheless, it provides accurate results for a large class of materials. All ab-initio methods have in common that they, just like the original Hamiltonian, require only the number of particles and the types of nuclei as their input parameters.

In the model Hamiltonian approach, one aims to write down a simplified Hamiltonian that captures the physics one expects to be dominant in a given system. The model Hamiltonians are easier to solve than the full Hamiltonian, but they usually introduce new adjustable parameters. For realistic calculations, those parameters have to be obtained from somewhere. They can be adjusted using experimental results or calculated by ab-initio methods. Another caveat is that it is not always clear how the choice for a model Hamiltonian can be justified. Because model Hamiltonians emphasize certain physical effects and neglect certain others, their results indicate how important those effects are for the specific behavior studied by the model. They can be employed to answer questions of a general nature. For example, a model of localized interacting spins can be used to study under which circumstances, such a system may be ferromagnetic.

Both approaches have their advantages and disadvantages. While the ab-initio methods can give accurate numerical results for real materials, they still do not provide general explanations. Conversely, model Hamiltonians do provide means of explaining the observed phenomena, but these explanations remain unsatisfying when the free parameters have to be fitted to experimental data. With ab-initio methods becoming more accurate and computational methods for model Hamiltonians becoming more efficient, the current trend is to combine the advantages of both approaches: Ab-initio methods are used to calculate from first principles the parameters that enter the model Hamiltonian.

1.2. The Hubbard Model

The Hubbard model, developed by Hubbard [7, 8, 9], Kanamori [10], and Gutzwiller [11], is the simplest many-body model including electron-electron interaction, which cannot be reduced to an effective single-particle model. Despite its apparent simplicity, it is applicable to a large number of systems and is used to model the physics of correlated electrons such as high-T_c superconductivity, Mott transitions in transition metal oxides, and organic conductors. The only exactly solvable case, however, is the ground state in one dimension [12]. An extraordinarily successful approximate method is Dynamical Mean Field Theory (DMFT) [13], which maps the lattice problem onto a single-site impurity problem that has to be solved self-consistently. For a comprehensive review, see [14].

In the following, we show how the Hubbard model can, at least partially, be derived from first principles. It arises from the tight binding approximation (TBA), so we begin with an introduction to this important method.

1.2.1. The Tight-Binding Approximation

In a first step, the nuclear and electronic degrees of freedom in the full Hamiltonian (1.2) are separated. In the Born-Oppenheimer approximation [15], it is observed that the nuclei are much slower than the electrons because of their much larger mass. When treating the electrons, we can therefore assume the nuclei to remain in place. The core coordinates, $\mathbf{R}_\alpha$ can thus be viewed as fixed parameters instead of quantum mechanical operators. The electronic ground state energy as a function of the core coordinates then gives an effective potential for the nuclei. Thus, we want to solve the electronic Hamiltonian

$$H_e = -\sum_{j=1}^{N_e} \frac{\mathbf{p}_j^2}{2m} - \sum_{j=1}^{N_n}\sum_{\alpha=1}^{N_n} \frac{Z_\alpha e^2}{|\mathbf{r}_j - \mathbf{R}_\alpha|} + \sum_{j<k}^{N_e} \frac{e^2}{|\mathbf{r}_j - \mathbf{r}_k|}. \tag{1.3}$$

For the coordinates of the nuclei, we assume that they form a regular crystal lattice as we are not concerned with disordered solids in this thesis.

The tight-binding approximation can be introduced in two different ways, and we present both. A general introduction is given in Chapter 10 of [16]. The first approach starts from the atomic limit. Suppose the lattice constant is large. Then, the lattice can be treated as a collection of independent atoms so that the electrons are best described by the single particle atomic wavefunctions centered at the atoms. If the $\psi_n(\mathbf{r})$ are a complete set of eigenstates of the atomic Hamiltonian at the origin, the corresponding set of eigenstates for the atom at $\mathbf{R}_i$ is $\psi_n(\mathbf{r} - \mathbf{R}_i)$. To obtain a wavefunction that is compatible with the lattice periodicity, we set

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}_i} \mathrm{e}^{i\mathbf{k}\cdot\mathbf{R}_i} \psi_n(\mathbf{r} - \mathbf{R}_i) \tag{1.4}$$

where $\mathbf{k}$ ranges over the first Brillouin zone. The resulting wavefunctions $\psi_{n\mathbf{k}}$ retain their atomic character while satisfying the Bloch condition. However, using the atomic

eigenfunctions ψ_n in (1.4) becomes problematic when the lattice constant is reduced and the atomic limit no longer applies. The main approximation in the TBA is to assume that the we can just replace the ψ_n in (1.4) by a new atomic wavefunction $\phi(\mathbf{r})$ that is the linear combination of a small number of localized atomic wavefunctions,

$$\phi_\mu(\mathbf{r}) = \sum_m b_{\mu m}\psi_m(\mathbf{r}) \tag{1.5}$$

which are then used to build the full electronic wavefunctions

$$\psi_{\mu\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{N}}\sum_{\mathbf{R}_i} \mathrm{e}^{i\mathbf{k}\cdot\mathbf{R}_i}\phi_\mu(\mathbf{r}-\mathbf{R}_i). \tag{1.6}$$

The index μ is called orbital index. This method is also called LCAO, for Linear Combination of Atomic Orbitals. One can derive an eigenvalue equation for the $b_{\mu m}$ by inserting (1.6) and (1.5) into the electronic Hamiltonian, but we will now move on to another way of arriving at suitable atomic orbitals $\phi(\mathbf{r})$ with some advantageous properties.

Suppose that part of the electron-electron interaction can be incorporated into an effective single-particle Hamiltonian, h^0_{eff}. From its solution, we obtain the Bloch states $|\phi_{\mathbf{k}\mu}\rangle$ and the band energies $\varepsilon_{\mathbf{k}\mu}$. From the Bloch waves, we obtain the *Wannier functions* [17] by means of a Fourier transformation:

$$|\psi_{i\mu}\rangle = \frac{1}{\sqrt{N}}\sum_{\mathbf{k}} \mathrm{e}^{-i\mathbf{k}\cdot\mathbf{R_i}}\,|\phi_{\mathbf{k}\mu}\rangle\,. \tag{1.7}$$

As the Bloch waves are orthonormal, so are the Wannier functions. The matrix elements of h^0_{eff} in terms of them are given by

$$\langle\psi_{i\mu}|h^0_{\text{eff}}|\psi_{j\nu}\rangle = \frac{1}{N}\sum_{\mathbf{k}\mathbf{k}'} \mathrm{e}^{i\mathbf{k}\cdot\mathbf{R}_i}\mathrm{e}^{-i\mathbf{k}'\cdot\mathbf{R}_j}\,\langle\phi_{\mathbf{k}\mu}|h^0_{\text{eff}}|\phi_{\mathbf{k}'\nu}\rangle = \frac{1}{N}\sum_{\mathbf{k}} \mathrm{e}^{i\mathbf{k}\cdot(\mathbf{R}_i-\mathbf{R}_j)}\delta_{\mu\nu}\varepsilon_{\mathbf{k}\mu}, \tag{1.8}$$

so they can be calculated by Fourier transforming the band energies. We define

$$\begin{aligned} \varepsilon_{i\mu} &= \langle\psi_{i\mu}|h^0_{\text{eff}}|\psi_{i\mu}\rangle \\ t_{ij\mu} &= \langle\psi_{i\mu}|h^0_{\text{eff}}|\psi_{j\mu}\rangle\,, \quad i\neq j \end{aligned} \tag{1.9}$$

with on-site energies $\varepsilon_{i\mu}$ and hopping amplitudes $t_{ij\mu}$ We note that for Wannier functions, $t_{ij\mu\nu} = t_{ij\mu}\delta_{\mu\nu}$, i.e., there is no intra-band hopping. In most cases, the Wannier functions are exponentially localized. That is, $|\psi_{i\mu}\rangle$ decays exponentially with $|\mathbf{r}-\mathbf{R}_i|$. This has, however, not been proven for the general case. A popular method is to choose the phase of the Bloch waves to obtain maximally localized Wannier functions [18]. Some results regarding localization properties of Wannier functions are given in [19, 20].

Both methods discussed so far finally yield orbitals $\psi_{i\mu}$ that are localized at the atoms. This allows us to restrict the $t_{ij\mu}$ to neighboring atoms. To arrive at a many-body formalism, we now introduce creation and annihilation operators $c^{(\dagger)}_{i\mu}$ for the Wannier basis. In second quantized form, the effective single particle Hamiltonian reads

$$H^0_{\text{eff}} = \sum_{i\mu,\sigma}\varepsilon_{i\mu}c^\dagger_{i\mu,\sigma}c_{i\mu,\sigma} - \sum_{i\neq j,\mu,\sigma} t_{ij\mu}c^\dagger_{i\mu,\sigma}c_{j\mu,\sigma}. \tag{1.10}$$

This is the tight-binding Hamiltonian. In most cases, hopping is restricted to nearest or next-nearest neighbors. The resulting Hamiltonian is a sparse matrix of low dimension and can therefore be diagonalized efficiently. However, the interaction is only accounted for in the effective single-particle potential $v_{\text{eff}}(\mathbf{r})$. We now move beyond the simple TBA and treat the interaction as a proper pair interaction.

1.2.2. Interaction in the TBA

Revisiting H_e, the electron-electron interaction can be partially incorporated into an effective single-particle potential [6]. To this end, we note that the expectation value of the electron-electron interaction can be written as a functional of the electron-pair density,

$$E_{\text{int}} = \left\langle \Psi \left| \sum_{j<k}^{N_e} \frac{e^2}{|\mathbf{r}_j - \mathbf{r}_k|} \right| \Psi \right\rangle = \int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'\, \rho_2(\mathbf{r}, \mathbf{r}') \frac{e^2}{|\mathbf{r} - \mathbf{r}'|}, \tag{1.11}$$

where the pair density ρ_2 is the expectation value of $\psi^\dagger(\mathbf{r})\psi\ (\mathbf{r})\psi^\dagger(\mathbf{r}')\psi\ (\mathbf{r}')$. We can, without loss of generality, write the pair density as

$$\rho_2(\mathbf{r}, \mathbf{r}') = \rho(\mathbf{r})\rho(\mathbf{r}')g(\mathbf{r}, \mathbf{r}') \tag{1.12}$$

with $g(\mathbf{r}, \mathbf{r}')$ being the *pair correlation function*. It is proportional to the probability of finding an electron at position $\mathbf{r}$ when we already know that there is an electron at position $\mathbf{r}'$. We rewrite ρ_2 even further, obtaining

$$\rho_2(\mathbf{r}, \mathbf{r}') = \rho(\mathbf{r})\rho(\mathbf{r}') + \rho(\mathbf{r})\rho(\mathbf{r}')(g(\mathbf{r}, \mathbf{r}') - 1). \tag{1.13}$$

Inserting (1.13) into (1.11) allows us to split the integral into

$$E_{\text{int}} = E_H + E_{xc} = e^2 \int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + e^2 \int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'\, \rho(\mathbf{r})\rho(\mathbf{r}') \frac{g(\mathbf{r}, \mathbf{r}') - 1}{|\mathbf{r} - \mathbf{r}'|}. \tag{1.14}$$

E_H is the Hartree term arising from the density-density interaction of the electrons, while E_{xc} contains the complicated exchange and correlation effects. In DFT, both of these energies are used to derive an effective single particle potential. From E_H, we obtain

$$v_{\text{eff}}(\mathbf{r}, \rho) = e^2 \int \mathrm{d}\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \tag{1.15}$$

which is the mean field of all electrons and therefore a functional of the electron density. From E_{xc}, we obtain

$$v_{xc}\mathbf{r}, \rho = \frac{\delta E_{xc}}{\delta \rho}. \tag{1.16}$$

To obtain a closed expression, however, one had to calculate the pair correlation function $g(\mathbf{r}, \mathbf{r}')$, which amounts to solving the many-body problem. In DFT, one approximates g to obtain an effective single-particle potential. We, in contrast, retain the exchange part as a full two-particle interaction. Nevertheless, we will make use of some of the properties

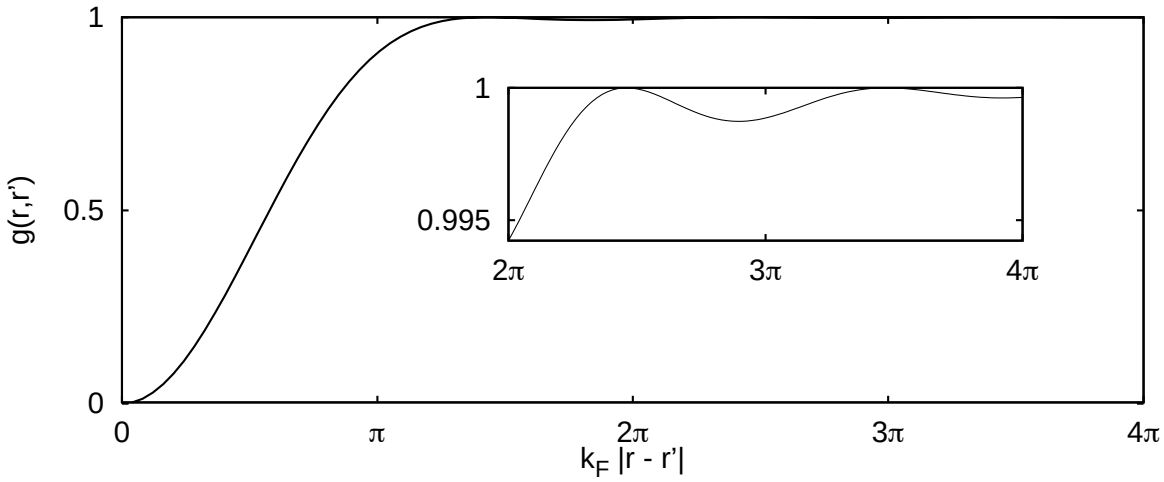


Figure 1.2.: Pair correlation function of the non-interacting homogeneous electron gas. It rapidly approaches 1, with oscillations on the scale of the Fermi wavenumber k_F. The part from 0 to π is called the exchange correlation hole, which is due to the Pauli exclusion principle.

of the pair correlation function g. An important observation is that $g(\mathbf{r}, \mathbf{r}') \to 1$ for $|\mathbf{r}, \mathbf{r}'| \gg 1$, as we expect correlations to be negligible for large distances. For the non-interacting electron gas, for example, the pair correlation function is shown in Fig. 1.2. Consequently, we expect the quantity $g(\mathbf{r}, \mathbf{r}') - 1$ to be local, as it approaches zero for large values of $|\mathbf{r} - \mathbf{r}'|$. This will be exploited later in our treatment of the interaction. For now, let us return to the single particle potential. We can write down an effective single particle Hamiltonian,

$$\begin{aligned} H^0_{\text{eff}} &= \sum_i h^0_{\text{eff}}(\nabla_i, \mathbf{r}_i) \\ h^0_{\text{eff}} &= h^0(\nabla, \mathbf{r}) + v_{\text{eff}}(\mathbf{r}, \rho) \\ v_{\text{eff}}(\mathbf{r}, \rho) &= e^2 \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \end{aligned} \tag{1.17}$$

The single-particle Hamiltonian h^0_{eff} has to be solved self-consistently, because the effective potential depends on the electron density, which in turn is determined by the solutions of h^0_{eff}. Note, however, that the density $\varrho(\mathbf{r})$ obtained from h^0_{eff} is not necessarily the true electron density. From its solution $|\phi_{\mathbf{k}\mu}\rangle$, $\varepsilon_{\mathbf{k}\mu}$, we can then obtain our Wannier functions (1.7).

Let us now write the interaction part of H_e in the Wannier basis. First, note that H_{int} can be written as

$$H_{\text{int}} = \frac{1}{2} \sum_{i \neq j}^{N_e} v^{\text{el-el}}(\mathbf{r}_i, \mathbf{r}_j), \qquad \text{with } v^{\text{el-el}}(\mathbf{r}, \mathbf{r}') = \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} \tag{1.18}$$

In the effective single-particle Hamiltonian (1.17), we already account for part of the interaction, so the residual interaction is

$$\bar{v}(\mathbf{r},\mathbf{r}') = v^{\mathrm{el-el}}(\mathbf{r},\mathbf{r}') - \frac{v_{\mathrm{eff}}(\mathbf{r},\rho) + v_{\mathrm{eff}}(\mathbf{r}',\rho)}{N_e}. \tag{1.19}$$

The full Hamiltonian reads

$$H = H^0_{\mathrm{eff}} + V^{\mathrm{el-el}}_{\mathrm{eff}} \tag{1.20}$$

where the interacting part, in second quantization, is

$$V^{\mathrm{el-el}}_{\mathrm{eff}} = \sum_{\substack{i\alpha,j\beta,k\gamma,l\delta \\ \sigma,\sigma'}} W^{\alpha\beta\gamma\delta}_{ijkl}\, c^\dagger_{i\alpha,\sigma} c^\dagger_{j\beta,\sigma'} c_{k\gamma,\sigma'} c_{l\delta,\sigma} \tag{1.21}$$

where the matrix elements $W^{\alpha\beta\gamma\delta}_{ijkl}$ are best evaluated using the real space representation of the Wannier functions:

$$W^{\alpha\beta\gamma\delta}_{ijkl} = \int \mathrm{d}\mathbf{r} \int \mathrm{d}\mathbf{r}'\, \bar{v}(\mathbf{r},\mathbf{r}') \psi^*_{i\alpha}(\mathbf{r}) \psi^*_{j\beta}(\mathbf{r}') \psi_{k\gamma}(\mathbf{r}') \psi_{l\delta}(\mathbf{r}). \tag{1.22}$$

Many of these terms become negligibly small. The exponential localization of the Wannier functions makes terms with $i \neq j \neq k \neq l$ small while the correlation hole is responsible for terms with $i = j \neq k = l$. For density-density terms ($i = l, j = k$). In the Hubbard model, one makes the drastic approximation that a nonzero matrix element is obtained only for $i\alpha = j\beta = k\gamma = l\delta$, resulting in

$$\sum_{i\alpha\sigma\sigma'} W^{\alpha\alpha\alpha\alpha}_{iiii} c^\dagger_{i\alpha,\sigma} c^\dagger_{i\alpha,\sigma'} c_{i\alpha,\sigma'} c_{i\alpha,\sigma} = \sum_\alpha U_\alpha \sum_i n_{i\alpha\uparrow} n_{i\alpha\downarrow}. \tag{1.23}$$

Alternatively, the direct interaction between different orbitals centered at the same site can be combined so that the sum over α in (1.23) is dropped and we have

$$U \sum_i n_{i\uparrow} n_{i\downarrow}, \qquad \text{with } n_{i\sigma} = \sum_\alpha n_{i\alpha\sigma}. \tag{1.24}$$

There are two types of processes ignored in (1.21). First, there are the terms $V_{i\alpha,j\beta} := W^{\alpha\beta\beta\alpha}_{ijji}$, corresponding to *direct interaction*. These processes can be accounted for by additional density-density terms

$$\sum_{i\alpha\neq j\beta} V_{i\alpha,j\beta} n_{i\alpha} n_{j\beta}, \tag{1.25}$$

which we will include in our extended Hubbard model that is introduced in Chapter 3. Second, there are *exchange terms* of the form W_{ijij}. Although they give rise to interesting effects [21], they are not treated here except for the Pauli principle, which leads to the cancellation of terms of the form $n_{i\uparrow} n_{i\uparrow}$.

Bringing everything together, the Hubbard Hamiltonian reads

$$H = \sum_{i\mu,\sigma} \varepsilon_{i\mu} c^\dagger_{i\mu,\sigma} c_{i\mu,\sigma} - \sum_{i\neq j,\mu,\sigma} t_{ij\mu} c^\dagger_{i\mu,\sigma} c_{j\mu,\sigma}. + U \sum_i n_{i\uparrow} n_{i\downarrow}. \tag{1.26}$$

We can go back to the Bloch basis by Fourier-transforming again. This yields

$$H = \sum_{\mathbf{k}\mu,\sigma} \varepsilon_{\mathbf{k}\mu\sigma} c^\dagger_{\mathbf{k}\mu,\sigma} c_{\mathbf{k}\mu,\sigma} + \frac{U}{N} \sum_{\mathbf{k},\mathbf{k}',\mathbf{q},\mu} c^\dagger_{\mathbf{k}\mu,\uparrow} c_{\mathbf{k}-\mathbf{q},\mu,\uparrow} c^\dagger_{\mathbf{k}'\mu,\downarrow} c_{\mathbf{k}'+\mathbf{q},\mu,\downarrow} \tag{1.27}$$

1.2.3. The Hubbard U

In the process of deriving a simplified model Hamiltonian, one desires to keep only few bands of interest. For such a reduced, effective system, however, it is all but clear how the interaction strength U, which is also called the Hubbard U, can be obtained from first principles. Calculating U via (1.22) only gives the so-called *bare* interaction. By neglecting all the other bands, the interaction is renormalized due to screening effects. Often, the Hubbard U is adjusted until the results of the model agree with experiments, but this is unsatisfactory as it destroys the predictive power of the method. There are also some approaches for calculating the Hubbard U from first principles, such as LDA+U [22], c-RPA [23], and GW@LDA+U [24].

Another problem with the Hubbard U is that interactions are already partially taken into account by the effective single-particle potential. This *double counting* [25] has to be corrected, but how this is done in a rigorous way remains unknown. The problem in DFT is that the exchange correlation energy, in contrast to the Hartree term, cannot be separated into separate parts for the different bands.

1.3. Motivation

The intent of this thesis is to study the nature of the Hubbard parameters, t and U. In particular, we are interested in the influence of additional bands on the Hubbard parameters of a single-band model. We propose a multiband model system that can be solved exactly using the new supercomputers at *Forschungszentrum Jülich*. Studying the exact solution to this model, we aim to understand how the Hubbard parameters have to be adjusted when the model is reduced to an effective single-band model.

Chapter 2 explains how we characterize and numerically solve many-body electron systems using the computing facilities at Jülich. Chapter 3 introduces and discusses our multiband model. Screening effects are investigated in Chapter 4, with a phenomenological approach presented in 4.1 and a systematic approach in given in 4.2. We take a closer look at one of the existing methods for computing a Hubbard U from first principles, c-RPA, in Chapter 5, before we summarize our findings in Chapter 6.

2. Methods

The complexity of many-body physics is reflected in the computational effort required to tackle it. Even for a simple Hubbard model the Hilbert space grows exponentially with the number of lattice sites, rendering any direct attempt at diagonalization useless. Thus, if we ever are to arrive at a computational solution for a problem we must make explicit use of its underlying structure. An example for this foundational principle of computational science is the treatment of noninteracting electrons. Their Hilbert space is just as large as the one for interacting electrons, but there is a trivial factorization of the many-body Hamiltonian, reducing the problem to a single electron problem.

In this chapter, we will explain how the structure of the Hubbard model can be exploited to allow for efficient treatment on parallel and massively parallel machines. In section 2.1 we explain the mathematical foundation of the Lanczos method, an efficient iterative method for diagonalization. Section 2.2 introduces single-particle Green's functions, which are an important tool in many-body physics. The final section, 2.3, explains how the Lanczos method is implemented on modern parallel computer architectures.

2.1. The Lanczos Algorithm

The exponential growth of the Hilbert space one encounters in many-body problems makes the direct diagonalization of Hamiltonians in matrix form impossible; for most algorithms, the space and time complexity are $\mathcal{O}(n^2)$ and $\mathcal{O}(n^3)$ [26], where n is the dimension of the Hilbert space. An example is given in Tab. 2.1. As direct diagonalization requires the storage of the entire matrix H, whose size is $(\dim \mathcal{H})^2$, the memory requirements quickly exceed the capabilities of even the largest supercomputers. Therefore, we resort to an iterative method, the Lanczos algorithm [27], which allows us to compute an approximation to the ground state relying solely on matrix-vector products and other elementary operations. Even here, the limiting factor for the treatment of large systems is the memory, not computation time. We will first describe the mathematical ideas behind the algorithm and then explain how the matrix-vector products can be computed with high efficiency on modern multiprocessor architectures.

2.1.1. Ground State

At heart, the Lanczos method is a variational approach; the matrix whose ground state is to be calculated is diagonalized in a special subspace. The ground state energy obtained

Table 2.1.: Dimension of the Hilbert spaces for a system with L single-particle states at half filling ($N_\sigma = L/2$) and corresponding memory requirements.

L	dim of $\mathcal{H}_\sigma$	dim of $\mathcal{H}$	memory
2	2	4	
4	6	36	
6	20	400	
8	70	4900	
10	252	63504	500kB
12	924	853776	6.51MB
14	3432	11778624	89MB
16	12870	165636900	1264MB
18	48620	2363904400	17GB
20	184756	34134779536	254GB

this way is an upper bound to the true ground state energy:

$$\frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} \geq E_0. \tag{2.1}$$

What makes the Lanczos method so successful is the way the subspace is obtained. We will first describe the method and then explain the benefits of this choice.

Let $\mathcal{H} = \mathbb{C}^n$ be the Hilbert space of our system and let $H \in \mathbb{C}^{n\times n}$ be the Hamiltonian. For an arbitrary vector $\mathbf{b} \in \mathcal{H}$, we define its *Krylov space* [28] of size m as

$$\mathcal{K}_{b,m} = \left\langle \mathbf{b}, H\mathbf{b}, H^2\mathbf{b}, \ldots, H^{m-1}\mathbf{b} \right\rangle. \tag{2.2}$$

Many of the modern iterative eigenvector methods work with this subspace [29]. They all rely on a theorem regarding eigenpairs of invariant subspaces, so we will describe it here.

Subspace methods generally proceed as follows. An invariant subspace of $\mathcal{H}$ under H is computed, together with a suitable basis $Q = (\mathbf{q}_1, \ldots, \mathbf{q}_m)$. In this basis, the Hamiltonian has the matrix form T. This matrix is then diagonalized either directly or iteratively. As the generated subspace usually is much smaller than the full Hilbert space, this can be done efficiently. The diagonalization of T then yields eigenpairs from which we can obtain eigenpairs of H. If the subspace is not an invariant subspace, the resulting eigenpairs are only approximations to eigenpairs of H.

A linear subspace $M \subset \mathcal{H}$ is called invariant under H if for every $\mathbf{x} \in M$ we also have $H\mathbf{x} \in M$. When diagonalizing H in the subspace M, the resulting eigenpairs are also eigenpairs of H in the full space. More specific, let $Q = (\mathbf{q}_1, \ldots, \mathbf{q}_m)$ be the matrix of basis vectors for the invariant subspace M, that is, $Q \in \mathbb{C}^{n\times m}$. Because of the invariance under H, $H\mathbf{q}_i$ must be representable as a linear combination of the Q as well, so there must be a vector $\mathbf{t}_i \in \mathbb{C}^m$ with $H\mathbf{q}_i = Q\mathbf{t}_i$. With $T = (\mathbf{t}_1, \ldots, \mathbf{t}_m)$ we can write $HQ = QT$. Now let $(\mathbf{z}, \lambda)$ be an eigenpair of T. Then we have

$$\begin{aligned} T\mathbf{z} &= \lambda\mathbf{z} \\ \Leftrightarrow QT\mathbf{z} &= \lambda Q\mathbf{z}. \\ \Leftrightarrow HQ\mathbf{z} &= \lambda Q\mathbf{z} \end{aligned} \tag{2.3}$$

Thus, $Q\mathbf{z}$ is an eigenvector of H with eigenvalue λ. Note that $Q\mathbf{z}$ and $\mathbf{z}$ are the same vectors written in a different basis. If Q is unitary, we have an even stronger connection between H and T, because in this case we have $Q^\dagger HQ = T$. In other words, T is the matrix representation of H in the subspace basis,

$$T_{ij} = \langle \mathbf{q}_i | H | \mathbf{q}_j \rangle \,. \tag{2.4}$$

The preceding observations hold for all invariant subspaces, but it is not said how such a subspace together with a suitable basis Q is computed. In the power method, one sets $Q = (\mathbf{b}, H\mathbf{b}, \dots H^{m-1}\mathbf{b})$. This generates an invariant subspace as soon as $H^{m-1}\mathbf{b}$ is linearly dependent on its predecessors. When n is the dimension of the full Hilbert space, we obtain an invariant subspace after at most n iterations. The power method has, however, several drawbacks which are addressed in the literature [30, 31]. In the Lanczos method, each new vector $q_i = Hq_{i-1}$ in the Krylov sequence is orthogonalized with respect to its predecessors. In particular, we have

$$\begin{aligned} \mathbf{q}_{i+1} &= \frac{H\mathbf{q}_i - \alpha_i \mathbf{q}_i - \beta_{i-1}\mathbf{q}_{i-1}}{\beta_i}, \\ \text{with} \quad \alpha_i &= \langle q_i | H | q_i \rangle \,,\, \beta_i = ||H\mathbf{q}_i - \alpha_i \mathbf{q}_i - \beta_{i-1}\mathbf{q}_{i-1}||_2. \end{aligned} \tag{2.5}$$

The subspace generated is the same as in the power method, but the basis Q is numerically advantageous. In the numerous literature (for a very good summary, see, e.g., [30]), the following positive properties of this method are described.

- To compute a new Lanczos vector, we only need to store its two immediate predecessors (Eq. 2.5).

- The resulting matrix T is hermitian and tridiagonal,

$$T = \begin{pmatrix} \alpha_0 & \beta_1 & & & \\ \beta_1 & \alpha_1 & \beta_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \beta_{m-1} & \alpha_{m-1} & \beta_m \\ & & & \beta_m & \alpha_m \end{pmatrix}, \tag{2.6}$$

  allowing for efficient diagonalization techniques [26]. In fact, many eigenvector solvers first transform a given matrix into tridiagonal form.

- To determine the extremal eigenpairs of H, it is not necessary to wait until an invariant subspace is reached. It turns out that the extremal eigenpairs of T quickly converge to the extremal eigenpairs of the full Hamiltonian, provided the start vector $\mathbf{b}$ is not orthogonal to them. In this unfortunate case, we obtain an invariant subspace that does not contain the ground state at all. Results should therefore be checked for reproducibility using different start vectors.

The last two points are important for computational considerations. First, we only have to store a few of the large many-body vectors from the full Hilbert space. Second, we can stop the generation of new Lanczos vectors after a small number (up to 200) of iterations. This keeps the matrix T at a modest size, so that exact diagonalization with direct methods remains feasible. This extraordinary convergence rate of the Lanczos algorithm has been studied by Kaniel [32] and Paige [33]. Their proofs are too long to reproduce here, so we only give short motivation. In essence, the Lanczos vectors are those one would generate in a gradient descent (ascent) method for minimizing (maximizing)

$$\frac{\langle \mathbf{q}|H|\mathbf{q}\rangle}{\langle \mathbf{q}|\mathbf{q}\rangle}, \tag{2.7}$$

which is equivalent to finding the smallest (largest) eigenpair of H.

We now give a pseudo-code description of the Lanczos algorithm as used for the calculations in this thesis. It is taken from [31] and repeated here for easier reference. The algorithm works in two passes; in a first pass, the ground state energy and the matrix elements of T are calculated while a second pass then computes the ground state vector. The first pass is shown in algorithm 1. The function pmassign in line 2 performs the matrix-vector multiplication of H and $\mathbf{b}$ and adds the result to $\mathbf{q}$. This is the most time consuming part of the algorithm. We will later describe how this product can be computed efficiently. Lines 3 and 4 make use of basic linear algebra operations, namely dot and axpy (for a times x plus y). These can be either taken from BLAS[1] or programmed and parallelized by hand. In line 6 we first create the symmetric tridiagonal matrix whose diagonal elements are given by the α_i and whose off-diagonal elements are given by the β_i. Then, we compute its lowest eigenvalue using one of the methods from LAPACK[2]. The LAPACK and BLAS packages contain plenty of highly optimized routines for basic and advanced linear algebra purposes. Note that for the first pass we only need to store two many-body vectors at a time, $\mathbf{b}$ and $\mathbf{q}$. The steps necessary to compute $\mathbf{q}_{i+1}$ are carefully arranged for optimal numerical stability [34]. In the algorithm, there are three points at which the iteration can stop. First, in line 1 we limit the total number of iterations performed to put an upper limit to the size of the tridiagonal matrix T. Second, in line 5 we test if an invariant subspace was reached, which is indicated by very small values of β. For large Hilbert spaces, however, this is unlikely to happen. As we are interested in the ground state only, we therefore have a check in line 7 that stops the iteration once the change in the ground state energy between two steps of the iteration becomes negligibly small, indicating that it has converged.

The diagonalization of the matrix T gives us a ground state vector $\mathbf{y}$ in terms of the Lanczos basis. To obtain the ground state vector $\mathbf{r}$ in the full space, we need the Lanczos vectors, since $\mathbf{r} = Q\mathbf{y}$. In principle, one could have stored the Lanczos vectors generated in algorithm 1 to have them readily available after $\mathbf{y}$ is computed. However, this will require the storage of m many-body vectors, where m can be on the order of a few hundred. As pointed out at the beginning of this chapter, memory requirements are our main concern,

[1] http://www.netlib.org/blas/
[2] http://www.netlib.org/lapack/

Algorithm 1: Lanczos, First Pass

Input: start vector $\mathbf{b}$, scratch space $\mathbf{q}$, Hamiltonian H, maxiter, precision, breaking condition *cond*
Result: Ground state energy, matrix elements α_i, β_i

```
begin
    clear(q)
    j := 0
1   while j ≤ maxiter do
        if j ≠ 0 then
            b := −β_{j−1} b
            q := (1/β_{j−1}) q
            swap(b,q)
        end
2       pmassign(H,b,q)
3       α_j := ⟨b|q⟩ // dot
4       q := q − α_j b // axpy
        if j < maxiter − 1 then
            β_j := ||q||_2
5           if β_j ≤ precision then // invariant subspace reached
                break
            end
        end
        j := j + 1
        if cond = "precision" or j = maxiter then
            E_old := E_new
6           E_new := GroundStateEnergy(SymTridiagMat(α,β))
            if j > 0 then
7               if |(E_old − E_new)/E_new| < precision then // converged
                    break
                end
            end
        end
    end
    return E_new,α,β
end
```

so it is not feasible to store all the Lanczos vectors. Fortunately, there is a way out. In a second pass of the algorithm, we perform the steps to create the Lanczos vectors again, but this time we incrementally build their linear combination $\mathbf{r}$ according to $\mathbf{y}$. Basically, we exchange the memory storage problems for a longer computation time. This way, we need storage space for only three many-body vectors, but we have to create the Lanczos

vectors for a second time. Two many-body vectors are needed to compute the next Krylov vector, while the third one is used to build the ground state vector. The procedure is shown in algorithm 2. In line 1 we use one of the LAPACK routines to compute the ground state vector of the tridiagonal matrix T. In line 2 we use the elements of $\mathbf{y}$ to build the ground state vector in the full Hilbert space.

Algorithm 2: Lanczos, Second Pass

Input: start vector $\mathbf{b}$, scratch space $\mathbf{q}$, Hamiltonian H, tridiagonal matrix T
Result: Ground state vector $\mathbf{r}$
begin
1 $\quad$ $\mathbf{y} :=$ `GroundState(`T`)`
$\quad$ `clear(q),clear(r)`
$\quad$ **for** $j = 0 \ldots$ `length(y)`-1 **do**
$\quad\quad$ **if** $j \neq 0$ **then**
$\quad\quad\quad$ $\mathbf{b} := -\beta_{j-1}\mathbf{b}$
$\quad\quad\quad$ $\mathbf{q} := \frac{1}{\beta_{j-1}}\mathbf{q}$
$\quad\quad\quad$ `swap(b,q)`
$\quad\quad$ **end**
2 $\quad\quad$ $\mathbf{r} := \mathbf{r} + y_j\mathbf{b}$
$\quad\quad$ `pmassign(`H`,b,q)`
$\quad\quad$ $\mathbf{q} := \mathbf{q} - \alpha_j\mathbf{b}$
$\quad$ **end**
$\quad$ **return r**
end

We conclude this section with a discussion of the loss of orthogonality. An accessible treatment is found in [30]. The results are due to Paige [33, 35]. In theory, all Lanczos vectors $\mathbf{q}_i$ should be mutually orthogonal, but this property can get lost due to numerical errors. This happens in particular when the β_i are small, which is equivalent to saying that cancellation errors occur in the computation of the next Lanczos vector. As this non-orthogonality deteriorates the accuracy of the eigenvalues, this has to be rectified. The initial suggestion by Lanczos himself was to reorthogonalize each new Lanczos vector with respect to its predecessors. However, this is very costly and nullifies the other advantages of the Lanczos method. As the analysis of Paige shows, loss of orthogonality goes hand in hand with convergence, so in our case we just stop the iteration as soon as the β_i become too small. In addition, the ground state energy usually converges long before an invariant subspace is reached, so the iteration stops before orthogonality is lost. Readers interested in more sophisticated remedies to the problem of orthogonality are referred to the literature [35, 30, 36].

Block Lanczos

If the first p eigenpairs of H are required, the Lanczos algorithm as described in the previous section is not adequate, because the first p eigenpairs of T do not necessarily correspond to the first ones of H. Instead, a block matrix version of the algorithm is used [37]. The single Lanczos vector $\mathbf{q}_i$ is replaced by an entire block X_i of size $n \times p$. The resulting matrix T will be band-diagonal with bandwidth p:

$$T = \begin{pmatrix} M_1 & B_1^T & & \\ B_1 & \ddots & \ddots & \\ & \ddots & \ddots & B_{r-1}^T \\ & & B_{r-1} & M_r \end{pmatrix}, \tag{2.8}$$

where the B_i are upper triangular and the M_i are symmetric. The X_i are generated via

$$R_k = HX_k - X_k M_k - X_{k-1} B_{k-1}^T \tag{2.9}$$

$$X_{k+1} B_k = R_k \qquad (QR\text{-factorization of } R_k) \tag{2.10}$$

$$M_{k+1} = X_{k+1}^T H X_{k+1}, \tag{2.11}$$

which is the block matrix analogue to (2.5). Most properties of the single vector Lanczos method carry over to the block Lanczos method. For example, the X_k are mutually orthogonal provided the R_k have rank p. If the rank of R_k is smaller than p, then X_{k+1} can still be constructed so that $X_{k+1}^T X_i = 0$ for $1 \leq i \leq k$ [38]. The main treats of the block Lanczos algorithm are [30]:

- The p smallest (largest) eigenvalues of the matrix T are good approximations to the smallest (largest) eigenvalues of H [39]. The error bound improves with larger block size.
- The diagonalization of T has time complexity $\mathcal{O}(p^2)$.
- Degenerate eigenpairs are correctly identified when p is at least as large as the largest multiplicity of any eigenvalue.

The trade off between better error bounds and more costly iterations with larger p is discussed in [40].

2.1.2. Spectral Functions

In this section, we discuss how the Lanczos method can be used to gain information about the full spectrum of a Hamiltonian. We defer the discussion of the underlying physics to the next section and focus for now on the technical aspects.

Let us just say that, given an arbitrary operator $\hat{A}$ acting on our many-body states, we want to compute the quantity

$$I_A(\omega) = -\frac{1}{\pi} \Im \left\langle \psi_0 | \hat{A}^\dagger \left[\omega - (H - E_0) + i\eta \right]^{-1} \hat{A} \, | \psi_0 \right\rangle, \tag{2.12}$$

which we call *spectral function*. For actual calculations, we need a small, nonzero η. In the following derivation, however, we take the limit $\eta \to 0$. The origin of η will be explained in Sec. 2.2.1. If $\mathcal{X}$ denotes the Hilbert space the vector $\hat{A}\ |\psi_0\rangle$ lies in, we can transform eq. 2.12 into the *spectral representation*,

$$I_A(\omega) = \sum_n \left|\langle\psi_n^{\mathcal{X}}|\hat{A}|\psi_0\rangle\right|^2 \delta\left(\omega - (E_n^{\mathcal{X}} - E_0)\right), \tag{2.13}$$

where the $|\psi_n^{\mathcal{X}}\rangle$ denote a complete orthogonal eigenbasis of H in the Hilbert space $\mathcal{X}$. For finite η, the δ-function is broadened into a Lorenz peak. In order to compute $I_A(\omega)$, we need the following quantities.

1. The ground state energy E_0. This can be computed using the Lanczos method, as discussed in the last section.

2. The eigenvalues of H in the Hilbert space to which $\hat{A}\,|\psi_0\rangle$ belongs. They are obtained by using the Lanczos method for a third pass, but instead of using a random start vector, we use
$$|\phi_0\rangle = \frac{\hat{A}\,|\psi_0\rangle}{\sqrt{\langle\psi_0|A^\dagger\hat{A}|\psi_0\rangle}} \tag{2.14}$$

3. The *spectral weights* $c_n := |\,\langle\psi_n^{\mathcal{X}}|\hat{A}|\psi_0\rangle\,|^2$. These can be computed for all the eigenvectors that belong to the Krylov space generated from $|\phi_0\rangle$.

The third point requires more explanation than the first two. If the orthogonalized Lanczos basis obtained from $|\phi_0\rangle$ is $\{|\phi_n\rangle\}$, then the eigenvectors $|\psi_n^{\mathcal{X}}\rangle$ can be written as

$$|\psi_n^{\mathcal{X}}\rangle = \sum_j U_{nj}\,|\phi_j\rangle \tag{2.15}$$

for a unitary matrix U whose rows are the eigenvectors of T. Simple algebra then leads to

$$c_n = \left|\langle\psi_n^{\mathcal{X}}|\hat{A}|\psi_0\rangle\right|^2 = \langle\psi_0|\hat{A}^\dagger\hat{A}|\psi_0\rangle\,|U_{n0}|^2. \tag{2.16}$$

Hence, in order to compute the spectral weights we require the norm of $\hat{A}\,|\psi_0\rangle$ and the first element of each eigenvector of T. Since a full diagonalization of T is computationally feasible, we thus have an efficient way of computing the spectral weights. While the Lanczos method only yields a subset of the eigenvectors of the full system, it turns out that these are the ones with the highest spectral weights. A full discussion of this, in addition to an alternative way of computing the spectral weights, is found in a review by Dagotto [41].

Off-diagonal Spectral Function

In the definition of $I_A(\omega)$ (2.12), the same operator $\hat{A}$ is used on both sides of the residual. If we imagine this operator to be part of a set of operators, $\hat{A}_i$, we call the $I_{A_i}(\omega)$ the diagonal elements of the spectral function. The previous section explained how these diagonal elements can be computed. Now suppose we want to compute the off-diagonal elements of the spectral function,

$$I_A^{\alpha\beta}(\omega) = -\frac{1}{\pi}\Im\left\langle\psi_0\middle|\,\hat{A}_\alpha^\dagger\left[\omega-(H-E_0)+i\eta\right]^{-1}\hat{A}_\beta\,\middle|\psi_0\right\rangle. \tag{2.17}$$

This can be reduced to the task of computing two diagonal spectral functions. Let

$$\hat{A}_\pm^{\alpha\beta} = \frac{1}{\sqrt{2}}\left(\hat{A}_\alpha \pm \hat{A}_\beta\right). \tag{2.18}$$

Using Lanczos, we obtain the two diagonal spectral functions $I_{A_\pm}^{\alpha\beta}$. By explicitly inserting $(\hat{A}_\pm^{\alpha\beta})$ into 2.12 and multiplying out, we obtain

$$I_{A\pm} = \frac{1}{2}\left(I_A^{\alpha\alpha} + I_A^{\beta\beta} \pm (I_A^{\alpha\beta} + I_A^{\beta\alpha})\right). \tag{2.19}$$

From this, we obtain

$$I_A^{\alpha\beta} + I_A^{\beta\alpha} = I_{A+}^{\alpha\beta} - I_{A-}^{\alpha\beta}. \tag{2.20}$$

If we turn to the spectral representation,

$$I_A^{\alpha\beta} = \sum_n \langle\psi_0|A_\alpha^\dagger|\psi_n^{\mathcal{X}}\rangle\,\langle\psi_n^{\mathcal{X}}|A_\beta|\psi_0\rangle\,\delta\left(\omega-(E_n^{\mathcal{X}}-E_0)\right), \tag{2.21}$$

we see that $I_A^{\alpha\beta}$ is the complex conjugate of $I_A^{\beta\alpha}$, so that we have in (2.20)

$$\Re I_A^{\alpha\beta} = I_A^{\alpha\beta} = \frac{1}{2}\left(I_{A+}^{\alpha\beta} - I_{A-}^{\alpha\beta},\right) \tag{2.22}$$

where we used that the spectral function is, per definition, real valued. We see that two additional Lanczos passes, one with $A_+\,|\psi_0\rangle$ and one with $A_-\,|\psi_-\rangle$ as start vectors are needed. From the resulting diagonal spectral functions, the off-diagonal elements can be obtained.

2.2. Green's Functions

We will now give physical meaning to the quantities discussed in the previous section. An invaluable tool in studying complex many-body systems are dynamic response functions. Much can be learned from studying how a system responds to change. One such response function is the causal single-particle Green's function,

$$G_{i\sigma j\sigma'}(t,t') = -i\left\langle\psi_0\middle|\,\mathcal{T}\left(c_{i\sigma}(t)c_{j\sigma'}^\dagger(t')\right)\middle|\psi_0\right\rangle. \tag{2.23}$$

Here, $\mathcal{T}$ is the time ordering operator with

$$\mathcal{T}\left(\hat{a}(t)\hat{b}(t')\right) = \begin{cases} \hat{a}(t)\hat{b}(t) & t > t' \\ -\hat{b}(t)\hat{a}(t) & t < t' \end{cases} \tag{2.24}$$

for operators with fermionic commutation rules. The operators $c_{i\sigma}^{(\dagger)}$ are the creation and annihilation operators in the Heisenberg picture.

The causal single-particle Green's function measures how a particle (hole) added to the system propagates through it. For $t > t'$, it represents the probability of finding a particle in state $i\sigma$ at time t after a particle has been added in state $j\sigma'$ at time t'. Its significance stems from the fact that it can be calculated using a variety of methods [42, 43] and provides a wealth of information about the system. Most prominently, we can obtain the energies of the elementary excitations. In other words, the Green's function contains information about the full spectrum. We will now explain how this is done.

2.2.1. The Spectral Function

As we are interested in the system's energy spectrum, we will now move from time to frequency. First, note that for a time-invariant Hamiltonian, the causal Green's function only depends on the time difference $t - t'$. We can thus set t' to zero and are left with only one time variable. Second, for Hamiltonians that conserve spin, $G_{i\sigma j\sigma'}$ is zero for $\sigma \neq \sigma'$. Therefore, we drop the spin index in the next steps. We begin our transition to frequency – or energy – space by writing the time-dependent Green's function in the so called spectral representation, which is also called Lehmann representation.

$$G_{ij}(t) = \begin{cases} -i \sum_n \langle \psi_0 | c_i | \psi_n^{N+1} \rangle \langle \psi_n^{N+1} | c_j^\dagger | \psi_0 \rangle \, \mathrm{e}^{i(E_0 - E_n^{N+1})t} & t > 0 \\ i \sum_n \langle \psi_0 | c_j^\dagger | \psi_n^{N-1} \rangle \langle \psi_n^{N-1} | c_i | \psi_0 \rangle \, \mathrm{e}^{i(E_n^{N-1} - E_0)t} & t < 0 \end{cases} \tag{2.25}$$

Here, $(E_i^{N\pm1}, |\psi_i^{N\pm1}\rangle)$ are the eigenpairs of the Hamiltonian in the Hilbert space with one more (+) or less (-) electron.

Taking the Fourier transform of $G_{ij}(t)$ will give us the Green's function, $G_{ij}(\omega)$. The spectral representation of $G_{ij}(\omega)$ is

$$G_{ij}(\omega) = \sum_n \frac{\langle \psi_0 | c_i^\dagger | \psi_n^{N-1} \rangle \langle \psi_n^{N-1} | c_j | \psi_0 \rangle}{\omega + (E_n^{N-1} - E_0) - i\eta} + \sum_n \frac{\langle \psi_0 | c_i | \psi_n^{N+1} \rangle \langle \psi_n^{N+1} | c_j^\dagger | \psi_0 \rangle}{\omega - (E_n^{N+1} - E_0) + i\eta} \tag{2.26}$$

The term $i\eta$ is a small convergence factor that is due to the Fourier transform of the (implicit) step functions $\theta(t - t')$ in (2.25). It can be obtained in numerous ways. The most elementary approach is to introduce the convergence factor in the Fourier integral.

$$\int_{-\infty}^{\infty} \mathrm{e}^{(-i\omega-\eta)t} \theta(t) \mathrm{d}t = \int_0^{\infty} \mathrm{e}^{(-i\omega-\eta)t} \mathrm{d}t = \frac{i}{\omega - i\eta}. \tag{2.27}$$

One might be tempted to now take the limit $\eta \to 0$ and obtain i/ω, but this gives the wrong result for $\omega = 0$. Instead, we use a well known result from distribution theory to

obtain

$$\frac{i}{\omega - i\eta} = \mathcal{P}\left(\frac{1}{\omega}\right) + i\pi\delta(\omega), \tag{2.28}$$

where $\mathcal{P}$ denotes the principal value. Alternatively, the convergence factor arises naturally when the reverse Fourier transform is solved by contour integration. In (2.26), the first term describes the part of the spectral Green's function that is due to the extraction of an electron, while the second term describes the part due to the insertion of an electron. These processes are called photoemission (PE) and inverse photoemission (IPE), respectively. This is because they are similar to what happens when photons interact with the system. To describe real photoelectric effects, one would have to include matrix elements for the electron-photon interaction as well. From the spectral representation of $G_{ij}(\omega)$ we immediately see that there is a pole at $\pm(E_n^{N\pm1} - E_0)$ for each eigenstate $|\psi_n^{N\pm1}\rangle$ with the poles for PE (IPE) in the upper (lower) complex plane.

The spectral Green's function is a complex valued function. To obtain a real valued function, we define the *spectral function* as

$$A_{\alpha\beta}(\omega) = \frac{1}{\pi}\Im\left[G_{ij}^{\mathrm{PE}}(\omega)\right] - \frac{1}{\pi}\Im\left[G_{ij}^{\mathrm{IPE}}(\omega)\right]. \tag{2.29}$$

Applied to (2.26), we obtain

$$\begin{aligned} A_{\alpha\beta}(\omega) = \; & \frac{1}{\pi}\sum_n \langle\psi_0|c_i^\dagger|\psi_n^{N-1}\rangle\langle\psi_n^{N-1}|c_j|\psi_0\rangle \frac{\eta}{(\omega + (E_n^{N-1} - E_0))^2 + \eta^2} \\ & + \frac{1}{\pi}\sum_n \langle\psi_0|c_i|\psi_n^{N+1}\rangle\langle\psi_n^{N+1}|c_j^\dagger|\psi_0\rangle \frac{\eta}{(\omega - (E_n^{N+1} - E_0))^2 + \eta^2}. \end{aligned} \tag{2.30}$$

A term of the form $\frac{1}{\pi}\frac{\eta}{z^2+\eta^2}$ is a representation of $\delta(z)$ for $\eta \to 0^+$. When we take the limit, we get

$$\begin{aligned} A_{\alpha\beta}(\omega) = \; & \frac{1}{\pi}\sum_n \langle\psi_0|c_i^\dagger|\psi_n^{N-1}\rangle\langle\psi_n^{N-1}|c_j|\psi_0\rangle \,\delta(\omega + (E_n^{N-1} - E_0)) \\ & + \frac{1}{\pi}\sum_n \langle\psi_0|c_i|\psi_n^{N+1}\rangle\langle\psi_n^{N+1}|c_j^\dagger|\psi_0\rangle \,\delta(\omega - (E_n^{N+1} - E_0)). \end{aligned} \tag{2.31}$$

From the spectral function, we learn two things about the spectrum. The position of the peaks tells us the eigenenergies of the system. The height of the peaks tells us the weight of the corresponding eigenvectors in the spectral representation of $c_i^{(\dagger)}|\psi_0\rangle$.

We are now in possession of both a physical motivation and a mathematical framework for computing spectral functions. Let us now discuss how this is implemented.

2.3. Implementation

Up to now, the general framework of the Lanczos method was described. For a successful implementation, two important questions must be answered. First, what is the basis for the full Hilbert space, $\mathcal{H}$, and how is it encoded? Second, how can vector operations, especially the matrix-vector product, be computed efficiently? These two questions are intimately related and in this section we provide an answer.

2.3.1. Many-Body Basis States

To describe many-body electronic states, we first choose a suitable single particle basis $|\phi_i\rangle$ and define creation and annihilation operators $c^\dagger_{i\sigma}$ and $c_{i\sigma}$ for electrons with spin σ in state $|\phi_i\rangle$. The occupation number basis is then defined as the set of states

$$|\{n_{i\sigma}\}\rangle = \prod_{i\sigma} \left(c^\dagger_{i\sigma}\right)^{n_{i\sigma}} |0\rangle \tag{2.32}$$

where $0 \leq n_{i\sigma} \leq 1$ for fermions. These are Slater determinants; the resulting states are antisymmetric because of the fermionic commutation relations of the creation operators. Any many-body state can be written as

$$|\Psi\rangle = \sum_{\{n_{i\sigma}\}} \alpha(\{n_{i\sigma}\}) \, |\{n_{i\sigma}\}\rangle \tag{2.33}$$

with coefficients α that can generally be complex valued. However, all systems we study exhibit time reversal invariance. For those, real valued coefficients suffice. Because of the anti-commutation rules of Fermions, the ordering of the factors in the product (2.32) is important; it has to be chosen and fixed in advance, as changing the ordering introduces a Fermi sign. If the Hamiltonian preserves spin, a first step is to order the states by their spins:

$$|\Psi\rangle = \prod_i \left(c^\dagger_{i\downarrow}\right)^{n_{i\downarrow}} \prod_j \left(c^\dagger_{j\uparrow}\right)^{n_{j\uparrow}} |0\rangle \tag{2.34}$$

Apart from this internal ordering of the single particle states, there is the ordering of the many-body basis states $|\{n_{i\sigma}\}\rangle$. When stored in a vector $\mathbf{q}$, it tells us which state is addressed by a matrix element q_i. Therefore, this ordering determines the structure of the resulting Hamiltonian, as $H_{ij} = \langle \mathbf{q}_i | H | \mathbf{q}_j \rangle$. Different orderings result in different positions for the matrix elements. Although any ordering is permissible, the correct choice has tremendous influence on the performance of, for example, the matrix vector product. In contrast to the single particle ordering, we can and will change the many-body basis state ordering throughout calculations when it suits our needs. A detailed explanation of this is given in the next and following sections.

Basis Encoding

For electrons of a given spin, an occupation number basis state is completely determined by specifying for each single-particle state $|\phi_{i\sigma}\rangle$ whether it is empty or occupied. If we have L single-particle states, this information can be encoded as a binary string of length L. The i-th bit from the right indicates the occupation of state $|\phi_i\rangle$. For full many-body states with two electron spins as in (2.34), we have to specify two binary strings, one for each spin type. Throughout this thesis, we make the choice that spin-up states are applied before spin-down states. As an example, we have

$$|0101, 0100\rangle = c^\dagger_{3\downarrow} c^\dagger_{1\downarrow} c^\dagger_{3\uparrow} |0\rangle \, . \tag{2.35}$$

The binary strings can be read as integers and provide us with a tuple $(i_\downarrow, i_\uparrow)$ that completely determines the state. For the example above, we have the tuple $(5, 4)$. Such a tuple can be mapped to a single integer i via either of these two one-to-one transformations

$$\begin{aligned} i &= i_\downarrow + 2^L i_\uparrow \\ i &= i_\uparrow + 2^L i_\downarrow \end{aligned} \tag{2.36}$$

because $0 \leq i_\downarrow, i_\uparrow, < 2^L$. Thus, to represent a many-body state using the occupation number basis, we need an array of size 2^{2L} where the i-th element gives the coefficient for state $|i\rangle = |(i_\downarrow, i_\uparrow)\rangle$.

If the number of electrons is fixed, $N = N_\downarrow + N_\uparrow$, not all states of the form $(i_\downarrow, i_\uparrow)$ are valid but only those where the binary representation of i_σ has exactly N_σ bits set. A simple mapping as in (2.36) from states to array indices is then not possible. The solution is to enumerate all valid integers and store them in a lookup table, one for each spin type. We call these tables configuration arrays. An example for a system with 4 sites and 2 electrons is shown in Fig. 2.1. The i-th entry of the array for spin σ is the i-th integer that

0	0	0	0	0
1	0	0	0	1
2	0	0	1	0
3	0	0	1	1
4	0	1	0	0
5	0	1	0	1
6	0	1	1	0
7	0	1	1	1
8	1	0	0	0
9	1	0	0	1
10	1	0	1	0
11	1	0	1	1
12	1	1	0	0
13	1	1	0	1
14	1	1	1	0
15	1	1	1	1

$\Longrightarrow$

0	3
1	5
2	6
3	9
4	10
5	12

Figure 2.1.: Left: All configurations for a system with 4 sites. Highlighted are all configurations with $N_e = 2$ electrons, i.e., two bits set. Right: Resulting configuration array.

has exactly N_σ bits set. Let $T_\sigma(i_\sigma)$ denote the function that describes the lookup. Then a tuple $(i_\downarrow, i_\uparrow)$ stands for the state $|T_\downarrow(i_\downarrow), T_\uparrow(i_\uparrow)\rangle$. The indices i_σ run from 0 to $\dim \mathcal{H}_\sigma$. The dimension of the Hilbert space for a single spin type is given by

$$\dim \mathcal{H}_\sigma = \binom{L}{N_\sigma}. \tag{2.37}$$

We can now provide suitable mappings from the tuple indices to a single integer index similar to (2.36):

$$i = i_\downarrow + \dim \mathcal{H}_\downarrow \cdot i_\uparrow \tag{2.38}$$

$$\text{or} \qquad i = i_\uparrow + \dim \mathcal{H}_\uparrow \cdot i_\downarrow. \tag{2.39}$$

We can use either of these mappings to define an array that represents a many-body state with $N_\downarrow$ down- and $N_\uparrow$ up-electrons. The total size of that array is equal to the dimension of the full Hilbert space,

$$\dim \mathcal{H} = \dim \mathcal{H}_\downarrow \times \dim \mathcal{H}_\uparrow = \binom{L}{N_\downarrow}\binom{L}{N_\uparrow}. \tag{2.40}$$

Exemplary sizes of the $\mathcal{H}_\sigma$ and $\mathcal{H}$ as well as the corresponding computer memory required to store the many-body state are shown in Tab. 2.1. From this table, we see that while $\mathcal{H}_\sigma$ remains modestly sized so that the configuration array can be stored without complications, the full Hilbert space and the storage requirements for many-body states quickly exceed the available working memory on normal computers. This calls for parallel architectures with distributed memory, which are discussed in the next section.

If we are interested in the configuration described by a single integer i, we first invert the mapping (2.38) or (2.39) to obtain $(i_\downarrow, i_\uparrow)$. Then, we use the configuration arrays to look up the elements at positions i_σ. Table lookup needs only constant time ($\mathcal{O}(1)$), so looking up a configuration for an index is very efficient. If we are interested in the single integer that describes a given configuration $|s_\downarrow, s_\uparrow\rangle$, we have to find the i_σ with $T_\sigma(i_\sigma) = s_\sigma$. As the configuration array is sorted, this can be done by binary search in time $\mathcal{O}(\log_2 \dim \mathcal{H}_\sigma)$. It is slower than the normal lookup, but this lookup is needed only during setting up the many-body Hamiltonian, which is done only once for a given system. It is not needed in the matrix-vector product itself, which is the most demanding part of our computations. Therefore, we prefer this binary search over setting up a lookup table for the reverse direction, because that would require 2^L entries.

Choice of Basis Set

The basis state encoding introduced in the previous section works for any choice of single particle basis $|\phi_n\rangle$. This freedom should be used to find a basis that is best suited for a given problem. The models we study in this thesis are related to the Hubbard model. As we recall from the introduction, in the real-space representation 1.26 the interaction is diagonal, while in the k-space representation 1.27 the hopping is diagonal. In most cases, hopping is restricted to nearest or next-nearest neighbors. Thus, the kinetic energy part of the real-space Hamiltonian is sparse, i.e., only a small number of matrix elements are nonzero. In the k-space representation, the interaction part is *not* sparse as electrons with any momentum can interact with each other. As the interactions are expected to cause the most trouble, we prefer a basis where the interaction part itself is of a simple form. This, together with the sparsity of the hopping part, is the advantage of the real space basis, because it allows for very efficient (sparse) matrix multiplication. For a multi-band

Hubbard model, we pick a basis $|\phi_{i\nu}\rangle$ where i is the site-index and ν is the band-index. The states are first ordered by site and then by band-index.

2.3.2. The Hamiltonian

We separate the Hamiltonian into a kinetic and an interaction part,

$$H = H_t + H_U. \tag{2.41}$$

We chose the basis in a way that makes H_U diagonal. To store it, we therefore need an array of size $\dim \mathcal{H}$, the same as for the many-body vectors. If we were to store the kinetic part, H_t, as a full matrix, we would require n^2 matrix elements where n is the size of the full Hilbert space. For a system with 12 sites at half filling this would require memory on the terabyte scale. Fortunately, there is a better way.

An important observation regarding the many-body Hamiltonian of the Hubbard model and its relatives is that the electron spin is preserved. Thus, we can also factorize the kinetic Hamiltonian:

$$H_t = (H_{t,\downarrow} + \mathbb{1}_\uparrow) + (H_{t,\uparrow} + \mathbb{1}_\downarrow) \tag{2.42}$$

where $H_{t\sigma}$ acts on states in $\mathcal{H}_\sigma$ only. We have already seen that the Hilbert spaces $\mathcal{H}_\sigma$ are significantly smaller than the full Hilbert space. The next important observation is that most entries of $H_{t\sigma}$ will be zero when hopping is restricted to the (next) nearest neighbors, because then only a small number of transitions can occur. An upper bound to the number of nonzeros of each row is obtained by dropping the Fermi exclusion principle. In one dimension, we get $2N_\sigma$ possible hopping transitions for nearest neighbor hopping and $4N_\sigma$ for next nearest neighbor hopping where N_σ is the number of electrons of spin σ, which is exponentially smaller than the size of the Hilbert space.

A storage scheme that is well adjusted to this situation and also has a nice physical interpretation is the Ellpack-Itpack format [44, 26]. Suppose we have a square matrix A of size $n \times n$. Let m be the maximum number of nonzeros per row. Then the nonzeros of A are stored in an array of size $n \times m$ where the nonzeros of each row are written consecutively, with the zeros between them omitted. In a second array of the same size, the column corresponding to each nonzero is stored. Finally, a one-dimensional array of length n stores the number of nonzeros per row. Below is an example of this scheme.

$$\begin{pmatrix} \varepsilon & -t & 0 & 0 & 0 \\ -t & \varepsilon & -t & 0 & 0 \\ 0 & -t & \varepsilon & -t & 0 \\ 0 & 0 & -t & \varepsilon & -t \\ 0 & 0 & 0 & -t & \varepsilon \end{pmatrix} \Longrightarrow \underbrace{\begin{array}{|c|c|c|} \varepsilon & -t & \\ -t & \varepsilon & -t \\ -t & \varepsilon & -t \\ -t & \varepsilon & -t \\ -t & \varepsilon & \end{array}}_{\text{nonzeros}}, \quad \underbrace{\begin{array}{|c|c|c|} 1 & 2 & \\ 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & \end{array}}_{\text{columns}}, \quad \underbrace{\begin{array}{|c|} 2 \\ 3 \\ 3 \\ 3 \\ 2 \end{array}}_{\text{length}}. \tag{2.43}$$

The relation to the physical situation is as follows. If the full matrix has a nonzero element t in row i and column j there is a transition with amplitude t from configuration i to configuration j. Now, for each configuration i the second array stores those configurations

j that are reachable from i, while the first array stores the corresponding transition amplitude. Computing these tables for our many-body Hamiltonian is one of the few situations where the reverse lookup mentioned in Sec.2.3.1 is needed. For a configuration given in binary representation $|\mathbf{n}\rangle$, where n is a bit-vector, we first obtain those configurations $|\mathbf{n}'\rangle$ that are reachable from $|\mathbf{n}\rangle$ and then use the reverse lookup to obtain the integer index j corresponding to $|\mathbf{n}'\rangle$.

Algorithm 3: Matrix-vector product for the Ellpack-Itpack format

Input: Matrix A in Ellpack-Itpack format (nzs,$cols$,$nnzs$), vector $\mathbf{v}$
Result: Vector $\mathbf{q} = A\mathbf{v}$
begin
  `clear(q)`
  for $i = 1..$`rows(`nzs`)` **do**
    for $k = 1..nnzs[i]$ **do**
      $\mathbf{q}[i] = \mathbf{q}[i] + nzs[i,k] \cdot \mathbf{q}[cols[i,k]]$
    end
  end
  return q
end

The matrix-vector product of a matrix in Ellpack-Itpack format is shown in algorithm 3. We use it separately for each spin part of the kinetic part of the Hamiltonian.

Now that the storage format of the Hamiltonian is specified, we can compare the memory requirements of the Lanczos method to those of a full diagonalization method. For Lanczos, we have to store three many-body vectors and the Hamiltonian, which consists of the sparse kinetic parts and the diagonal interaction part. All in all, we have

$$\mathsf{memory}_{\mathsf{Lanczos}} = \mathcal{O}(\dim \mathcal{H}) + \mathcal{O}(\dim \mathcal{H}_{t,\sigma}) = \mathcal{O}(\dim \mathcal{H}) \tag{2.44}$$

$$\mathsf{memory}_{\mathsf{full}} = \mathcal{O}\left((\dim \mathcal{H})^2\right). \tag{2.45}$$

2.3.3. Parallelization

To successfully implement the Lanczos algorithm on supercomputers, we have to make use of their parallel architecture. First, because using more CPUs speeds up the computation, and more importantly, because for larger systems the memory requirements exceed the working memory of a single computation node. The two main types of parallel systems are shared memory and distributed memory systems, which we will discuss in this section.

Shared Memory Parallelization

In a shared memory system, multiple CPUs work in parallel while accessing the same working memory. This is already found in modern desktop computers where, as of now, up to four CPUs are used. When examining algorithm 3, we note that the calculations for

different elements of the resulting vector, i.e., for different i, are independent of each other and can be performed in any order. This allows us to distribute the task of calculating the result for each i among multiple processors. This is achieved via the OpenMP[3] package, where a single preprocessor directive is used to tell the compiler how the loop over the indices can be parallelized. The speedup obtained this way is almost perfect, i.e., doubling the number of processors results in half the computation time. Precise measurements for the supercomputer found at Jülich Research Center have been performed in [31].

Distributed Memory Parallelization

When the Hilbert space becomes too large for many-body vectors to be stored in the RAM of a single computer, we must resort to supercomputers with multiple computation nodes and distribute the vectors among the nodes. While this resolves our memory shortage, it introduces new challenges, as many vector operations now involve time-consuming communication between the nodes. The chief goal in designing algorithms for distributed memory systems is therefore to reduce the amount of necessary communication. The communication between individual nodes is achieved using the procedures defined by the Message Passing Interface (MPI) standard[4]. Common implementations are OpenMPI[5] and MPICH[6].

The sparse matrices of the $H_{t,\sigma}$ are small enough to be stored locally on each computational node. All many-body vectors and H_U, however, have to be split up and distributed among the nodes. When we use tuples $(i_\downarrow, i_\uparrow)$ to denote the vector elements, we can view each vector as a *matrix* that is mapped to a linear array using either (2.38) or (2.39). Let us, for definiteness, suppose that the former is used. That is, $i_\downarrow$ denotes the row index and $i_\uparrow$ the column index. Suppose that the matrix is stored in column-major ordering, i.e., its columns are stored in memory one after another.

$$\begin{pmatrix} (0,0) \\ (0,1) \\ (0,2) \\ (1,0) \\ (1,1) \\ (1,2) \\ (2,0) \\ (2,1) \\ (2,2) \end{pmatrix} \quad \Longleftrightarrow \quad \begin{pmatrix} (0,0) & (0,1) & (0,2) \\ (1,0) & (1,1) & (1,2) \\ (2,0) & (2,1) & (2,2) \end{pmatrix}. \tag{2.46}$$

In this case, all states with the same configuration of spin-up electrons but different configurations of spin-down electrons are *local* in memory, because $(i_\downarrow, i_\uparrow)$ and $(j_\downarrow, i_\uparrow)$ are at most $\dim \mathcal{H}_\downarrow$ positions apart, while states $(i_\downarrow, i_\uparrow)$ and $(i_\downarrow, j_\uparrow)$ are at least $\dim \mathcal{H}_\downarrow$

[3] http://openmp.org
[4] http://www.mpi-forum.org
[5] http://www.open-mpi.org
[6] http://www.mcs.anl.gov/research/projects/mpich2/

positions apart. When we distribute the columns of the matrix among the nodes, all vector elements $(\cdot, i_\uparrow)$ are therefore stored on the same node.

As $H_{t,\downarrow}$ only operates on the down-electrons, it can then be applied to each column of the matrix independently. It does not require vector elements from other columns and, consequently, does not require communication among the processors. In contrast, $H_{t,\uparrow}$ operates on the rows. Since those are distributed among the nodes, applying $H_{t,\uparrow}$ requires communication. In a naive implementation, the vector elements could be exchanged on demand using so-called one-sided communication. A performance analysis by Andreas Dolfen [31], however, revealed that communication becomes so costly that the use of more nodes actually slows down the computation. In his thesis, he developed an elegant solution. In the matrix picture, we made the arbitrary choice to have $i_\downarrow$ as the row index. This resulted in the down-states being local in memory. If we now *transpose* the matrix so that $i_\uparrow$ becomes the row index, the up-states become local instead, allowing to efficiently apply $H_{t,\uparrow}$. Although such a matrix transposition still requires communication, this can be done in a very efficient way that was developed in [31, 45] and will be repeated here for easier reference.

The task is to transpose the matrix Q representing a many-body vector $\mathbf{q}$ that is stored in block-column-distribution. For simplicity, let us first assume there are $n = rp$ columns where p is the number of processors. Thus, each processor stores r columns. Let us for now also assume that $\dim \mathcal{H}_\uparrow = \dim \mathcal{H}_\downarrow$, so there are also n rows in Q. If we partition the matrix into blocks of size $r \times r$, we can write Q as a $p \times p$ block-matrix where processor i contains all blocks $Q_{i,\cdot}$ with

$$Q_{ij} = \begin{pmatrix} q_{ri,rj} & \cdots & q_{ri,r(j+1)-1} \\ \vdots & & \vdots \\ q_{r(i+1)-1,rj} & \cdots & q_{r(i+1)-1,r(j+1)-1} \end{pmatrix}, \quad Q = \begin{pmatrix} Q_{00} & \cdots & Q_{0,r-1} \\ \vdots & & \vdots \\ Q_{r-1,0} & \cdots & Q_{r-1,r-1} \end{pmatrix}. \tag{2.47}$$

To transpose Q, we first transpose with respect to the blocks so that Q_{ij} changes place with Q_{ji}, and then we transpose each Q_{ij} individually. The former requires communication between each processor i with each other processor j, while the latter is a local operation. The amount of communication necessary for the transpose might seem daunting at first, but there already is an MPI routine for this type of collective communication, the function `MPI_Alltoall`, with C prototype

```
int MPI_Alltoall(void *sendbuf, int sendcount, MPI_Datatype sendtype,
                 void *recvbuf, int recvcount, MPI_Datatype recvtype,
                 MPI_Comm comm)
```

When a processor encounters this command, it waits for all other processors to arrive at that point during their execution of the program. Then, every processor i sends to processor j the `sendcount` elements at position $j \cdot$ `sendcount` from `sendbuf` and receives `recvcount` from processor j into position $i \cdot$ `recvcount` of `recvbuf`, as shown in Fig. 2.2. For the block-transform, we call `MPI_Alltoall` for each of the r columns on a processor. For $n = rp$, the `sendcount` and `recvcount` are set to r. In this case, the same buffer can be used for sending and receiving, i.e., the exchange of elements is *in place*. If n is not an integer multiple of p, we need a version of the `MPI_Alltoall` command that gives

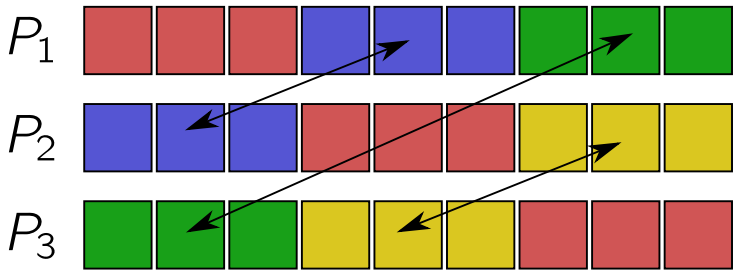


Figure 2.2.: Data sent and received by processors P_i when calling `MPI_Alltoall`

fined-grained control about the data sent and received. Such an advanced command is also necessary for non-square matrices and shall be described now.

We discuss the general case where Q is rectangular with n columns and m rows. Let

$$n = rp + r', \quad r' < r \tag{2.48}$$
$$m = tp + t', \quad t' < t. \tag{2.49}$$

In this case, the transpose operation becomes considerably more involved, as the number of elements sent by a processor differs from the number of elements received. However, the general idea of first exchanging blocks of data and subsequently performing local rearrangement remains. We will now outline the necessary steps. Before the transpose, processor i stores r_i columns. After the transpose, it might store a different number, k_i, of columns. These columns were rows before the transpose. Thus, we partition each column into blocks of lengths $k_0, k_1, \ldots, k_{p-1}$. Then, block i contains all elements destined for processor i. These blocks are distributed among the processors using `MPI_Alltoallv`, which is a generalization of `MPI_Alltoall`. It allows to specify the exact parts of the send and receive-buffers to use for the communication. The variables `sendcount` and `recvcount` are replaced by arrays `sendcounts` and `recvcount` where the i-th element specifies how many elements are sent (received) to (from) processor i. In addition, we have an array `senddispls` whose i-th element specifies at what position of `sendbuf` the data destined for processor i is to be found. The array `recvdispls` does the same for the received data. For the transpose, we iterate over all columns of a processor and set the arrays for the counts and displacements according to the rules. As the number of sent and received elements differs, we cannot perform the operation in-place. Instead, each processor allocates separate space `recvbuf` of size $k_i \cdot \dim \mathcal{H}_\uparrow$ so that this single buffer can hold all elements the processor will store after the transpose.

- `sendcounts`. As each processor P_i has its own version of the `sendcounts`-array, we define a matrix S_{ij} that stores the j-th element of the array on processor i. Suppose we are in the l-th iteration of the transpose algorithm. Then we have

$$S_{ij} = \begin{cases} k_j & r_i \geq l \\ 0 & \text{otherwise} \end{cases}. \tag{2.50}$$

That is, when there are columns left on processor i, we have to send as many elements to processor j as that processor will have columns after the transpose.

- `senddispls`. This array contains the positions of the blocks that are to be sent away. Because those are stored in memory consecutively, they are given by

$$\texttt{senddispls}[i] = \sum_{j=0}^{i-1} \texttt{sendcounts}[j]. \tag{2.51}$$

- `recvcounts`. Analogously to S_{ij} we define R_{ij} as the number of elements processor i receives from processor j. In principle, $R_{ij} = S_{ij}^T$, but we want to compute R_{ij} locally and S_{ij} is distributed among the processors. After the transpose, processor i will store k_i columns. This is the number of elements it will receive in each step from the other processors, provided the sending processor has any columns left at all. Thus, in step l, we have

$$R_{ij} = \begin{cases} k_i & r_j \geq l \\ 0 & \text{otherwise} \end{cases}. \tag{2.52}$$

- `recvdispls`. Setting the correct displacements for the receive buffer is not trivial. Unfortunately, it is not possible to set the values so that the elements are stored directly in the correct order. We have to collect the elements first and then bring them in the correct ordering. We describe the collection first. After iterating over all the columns, a processor P_i will have received elements that belong to an entire block of Q with size $k_i \times n$, i.e., an entire block row. Our aim is to store these elements in `recvbuf` so that, when seen as a matrix, `recvbuf` represents that block. To achieve this, we initialize

$$\texttt{recvdispls}[j] = j \cdot r \cdot k_j \tag{2.53}$$

so that elements received from processors j and $j+1$ are r columns apart. After each iteration, we update the displacement by

$$\texttt{recvdispls}[i] \mathrel{+}= \texttt{recvcounts}[i]. \tag{2.54}$$

Let `v` denote the array that holds the part of the many-body vector locally stored on processor P_i. When seen as a matrix, it has n rows and k_i columns. Its elements are given by

$$\texttt{v}[k_i \cdot n + l] = \texttt{recvbuf}[l \cdot k_i + k]. \tag{2.55}$$

This is just a local transpose, but because the blocks are rectangular rather than square, it cannot be implemented by simply swapping elements. Therefore, we use the additional `recvbuf` to first collect the elements and then shuffle them into the final vector via (2.55).

The entire process is sketched in Fig. 2.3. As discussed in [31], this transpose method shows excellent scaling behavior. We implemented both the OpenMP and MPI variants of the Lanczos algorithm and verified that it can be used to solve systems of up to 20 sites on Jugene, the new massively parallel supercomputer at Jülich.

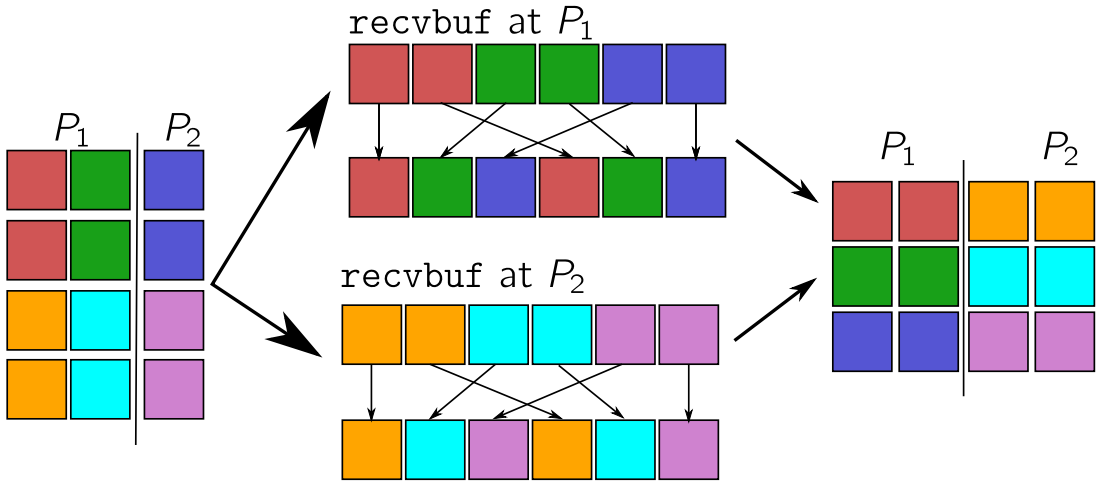


Figure 2.3.: Transpose of a rectangular matrix. The first step shows how the `MPI_Alltoallv` command distributes the columns. Colored elements are those sent to P_1. Elements of the same color are sent in the same iteration by the same processor. The next step shows how the receive buffer is shuffled into the vector `v`, whose matrix representation is shown in the last step.

Other vector operations. We conclude this section with a brief discussion of other vector operations that occur in the algorithms.

- Scaling a vector is done entirely local. If a node comprises many processors, it can be sped up using OpenMP directives. The same holds for adding two vectors.
- Computing the dot product of two vectors involves communication. Let $\mathbf{a}$ and $\mathbf{b}$ be the vectors and let $\mathbf{a}_i$ and $\mathbf{b}_i$ be their portions stored on processor i. First, each processor i computes $\langle \mathbf{a_i} | \mathbf{b_i} \rangle$. The results are then combined and re-distributed using

  ```
  int MPI_Allreduce( void* sendbuf, void *recvbuf, int count,
                     MPI_Datatype datatype, MPI_Op op, MPI_Comm comm )
  ```

  which takes the `count` elements from the `sendbuf` of each processor, combines them using the operator specified by `op` and sends the result back to each processor. In our case, the count is 1 and the operator is $+$. Because `MPI_Allreduce` is an important operation that occurs in many calculations, most architectures provide highly optimized implementations of it. Ideally, it scales with $\mathcal{O}(\log_2 p)$.

These operations are implemented in the standard way found in the literature [44].

3. Model

The main purpose of this project is to study the effect of screening on the parameters of the single-band Hubbard model. To this end, we design a model system with multiple electron bands which should satisfy the following requirements.

- The system must still be tractable with the computational methods introduced in Chapter 2. This is a question of Hilbert space dimension and Hamiltonian structure.
- The system should have a subsystem that can clearly be identified as a correlated band.
- The electron number in that subsystem should remain constant.

The second requirement simply states our aim of treating a single band that is embedded in a larger system using an effective single band theory. The third requirement is necessary for an effective single band theory to be viable. When we want to treat the subsystem as an effective single band, we do not want to allow for electrons to leave or enter the system. Although such systems can be treated within the formalism of open quantum systems [46], we want to keep the situation as simple as possible. As was explained in the introduction, many systems are treated by mapping them to a Hubbard model without further modifications.

Our choice of model is motivated by the typical band structure found in materials with a partially filled band. This band is the valence band; at zero temperature, all bands below are completely filled, while all bands above are completely empty. If the partially filled band is a d- or f-band, the electrons are strongly localized and thus lend themselves to the tight-binding model. As a consequence of this localization, correlation effects are of extreme importance. We therefore desire to describe them using the Hubbard model. To keep the model small, we do not want to include the other bands. The influence of the neighboring bands, however, is not negligible at all. Ideally, they merely renormalize the single-band Hubbard parameters, but they could just as well render the description in terms of a single-band model entirely useless. We aim to study both the renormalization and the breakdown of the effective single-band theory.

3.1. Three-band Hubbard model

Perturbation theory tells us that the most important bands to influence our band of interest are the immediate neighbors. For our studies, we therefore choose an extension of the Hubbard model where we have three orbitals at each lattice site. One for the core

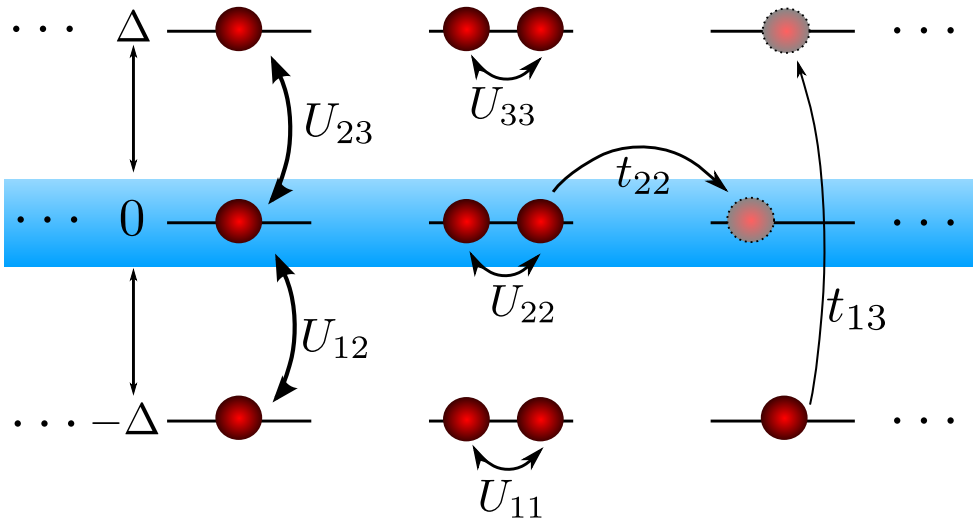


Figure 3.1.: Three-band Hubbard model.

states, one for the empty states and one for the partially filled correlated states. If not indicated otherwise, the system is at half filling. Fig. 3.1 shows a sketch of this three-band Hubbard model. At each lattice site we have three orbitals with energy separation Δ. In the central band, electrons can hop from site to site with hopping amplitude t_{22}. They can also move between the lower and the upper band with amplitude t_{13}. The electron-electron interaction is purely on-site and takes place only between electrons in the same orbital or between the central band electrons and the outer band electrons. We call the central band the *correlated* band. The electrons in the outer band are only weakly correlated, since the concentration of either electrons (in the upper band) or holes (in the lower band) is very low. To distinguish between electrons in the correlated band and the outer bands, we will call the former d-electrons and the latter r-electrons, similar to the decomposition introduced by F. Aryasetiawan [23].

The full Hamiltonian can be decomposed into H_d for the central band electrons and H_r for the other electrons. We can then write it as

$$H = H_d + H_r + H_{rd}. \tag{3.1}$$

For our three-band Hubbard model, the particular Hamiltonians are

$$H_d = -t_{22} \sum_{i,j:nn,\sigma} c^\dagger_{i,2\sigma} c_{j,2\sigma} + U_{22} \sum_i n_{i,2\uparrow} n_{i,2\downarrow} \tag{3.2}$$

$$H_r = -t_{13} \sum_{i\sigma} \left[c^\dagger_{i,1\sigma} c_{i,3\sigma} + c^\dagger_{i,3\sigma} c_{i,1\sigma} \right] + \sum_{i,\sigma} \Delta \left(n_{i,3\sigma} - n_{i,1\sigma} \right) + \sum_{\substack{i \\ \nu=1,3}} U_{\nu\nu} n_{i,\nu\uparrow} n_{i,\nu\downarrow} \tag{3.3}$$

$$H_{rd} = \sum_{\substack{i \\ \nu=1,3}} U_{2\nu} n_{i,2} n_{i,\nu}. \tag{3.4}$$

The indices 1, 2, and 3 refer to the three orbitals with 1 being the lowest. σ is the spin degree of freedom and nn denotes a sum over nearest neighbors. Let N_d denote the number of electrons in the central band and N_r the number of electrons in both outer bands. Then the Hamiltonian (3.1) conserves N_d and N_r. We therefore have two separate subsystems with Hamiltonian H_d for the d-electrons, H_r for the r-electrons, and H_{rd} for the Coulomb interaction between them. H_d itself is the familiar single-band Hubbard Hamiltonian. H_r can be interpreted as a collection of two-level atomic tight-binding Hamiltonians. H_{rd} contains the density-density interaction between the d- and r-electrons. It is implicitly introduced by t_{13}, which will be explained in Sec.3.2.

In the noninteracting ground state for half filling, the lower band is completely filled, the central band is half filled, and the upper band is completely empty. If now the inter-band interactions U_{12} and U_{23} are turned on, with $U_{12} > U_{23}$, we expect that the influence of the d-electrons will move electrons from the lower to the upper band. What we do not want to happen, however, is that electrons leave the central band due to the Coulomb repulsion. Instead, electrons added to a central orbital should stay in the central band, as shown in Fig. 3.2. To ensure that the system behaves as desired, we have to put some constraints on the Coulomb parameters. We will discuss this in the next section.

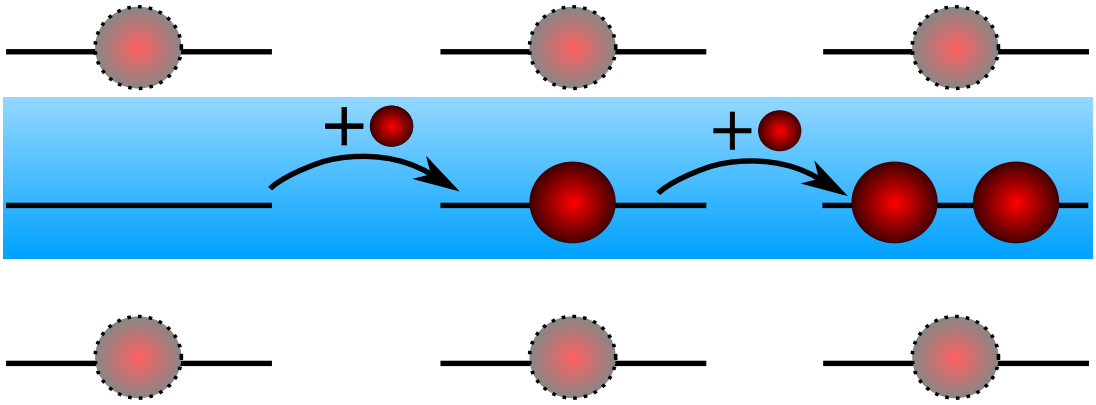


Figure 3.2.: Adding electrons to the central band. The parameters of the model have to be chosen in a way that ensures that the ground-state, depending on the number of electrons added, is always as in this sketch. That is, it should not be energetically favorable to have an electron enter or leave the central band

3.2. Parameters

Let us now review in detail the various parameters of our model to motivate their inclusion. We will begin with the most simple set of parameters and subsequently justify the introduction of more parameters.

In the simplest case, we have three bands with orbital energies E_1, E_2 and E_3. Using our freedom to move the zero level, we set $E_2 = 0$. To keep the model as simple as possible, we also demand that $E_3 = -E_1$. This leads to a single energy separation parameter Δ.

We also add the Hubbard parameters U_{22} and t_{22}, because these are the parameters whose renormalization we want to study. If we leave it at that, however, nothing interesting will happen, because then the outer bands have no influence on the central band. This makes the reduction to an effective single band model trivial.

Therefore, we also include the interaction energies U_{12} and U_{23}. This introduces a density-density interaction between the central band and the outer bands. Unfortunately, the desired screening behavior is still not achieved. In all the parameter sets we consider, we have $U_{12} > U_{23}$. The physical picture is that lower orbitals are more localized than higher orbitals and thus electrons in these orbitals experience a stronger Coulomb interaction. By comparing the total energy for each possible electron configuration, we find that a state where an electron has moved from the lower to the upper band once two electrons are in the central orbital can never be the ground state for reasonable parameter sets. For the proof, we look at the configurations shown in Fig. 3.3. Inserting the first

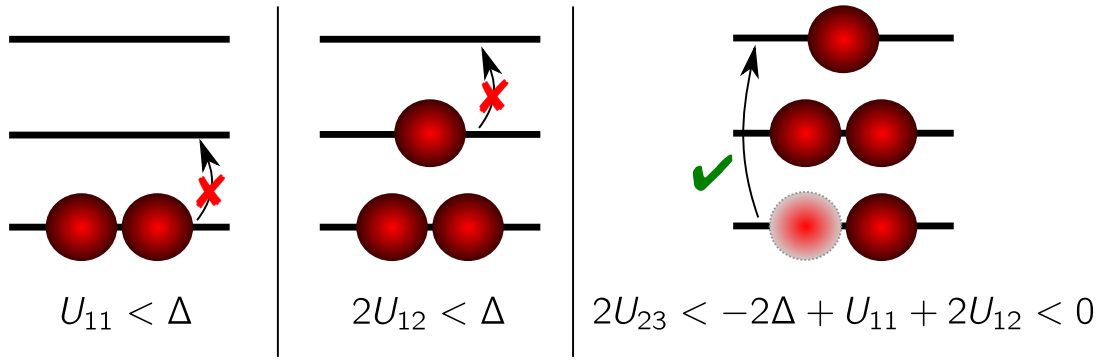


Figure 3.3.: Parameter inequalities ensuring the model system behaves as desired

two inequalities into the last one, we obtain $U_{23} < 0$. As we do not want to include an attractive interaction, there is no way to obtain the desired screening behavior. Besides, even if we were able to obtain it, another problem arises. Screening would always involve a full electron moving from the lower to the upper orbital. This is a very coarse type of screening. In particular, there is no difference in screening strength as it is always all or nothing. This can be partially improved by introducing hopping amplitudes t_{11} and t_{33}, allowing a screened electron to spread out over the entire lattice. However, the systems we are dealing with are small; a maximum of $N = 6$ sites can be treated. Hence, screening still involves at least one sixth of an electron, introducing severe finite size effects. What we want instead is a mechanism that allows for continuous screening.

As a remedy, we introduce a hybridization t_{13} between the upper and the lower orbitals. This hybridization introduces hopping and exchange terms in the new basis where the outer orbitals are orthogonal. One of these terms is a hopping term between the lower and the upper band that is proportional to the number of d-electrons on the site. A formal investigation of why the inclusion of a hybridization t_{13} is crucial to our system is given in Appendix A. This last step finally leads to the desired screening behavior where adding electrons to the central band moves electrons between the outer bands while keeping

the central band occupation fixed and where the occupation of the outer orbitals can be fractional.

Finally, let us look at U_{11} and U_{33}. Although they are not strictly necessary in order to obtain the desired model behavior, they provide additional control over the r-electrons. For all parameter sets we consider, we have $U_{11} > U_{33}$, because we assume the core electrons of the lower band to be more localized.

3.3. Basic Properties of the Model System

Before we analyze screening and renormalization effects, we will investigate the basic physics of our three-band model, to which we shall refer from now on as the *full system*. Our first goal is to find parameter regimes where the ground-state of the system shows the desired properties. We begin by looking at electron densities for a single site. Next, we present some results for the one-dimensional single-band Hubbard model. Finally, we look at the spectrum of the full system.

3.3.1. Atomic Limit

The atomic limit is obtained for $t_{22} = 0$. As the lattice sites become isolated, we can set N, the number of sites, to 1. We search for parameters that lead to the following behavior of the system.

- For $N_e = 2$, the central orbital is empty.
- For $N_e = 3$, the central orbital contains exactly one electron.
- For $N_e = 4$, the central orbital is doubly occupied. In addition, a transition in the r-electrons occurs, i.e., the electron density in the outer bands is different from that obtained for $N_e = 3$ and for $N_e = 2$.

In other words, electrons added to the system should always go into the d-band, and adding them should result in a change of the occupation in the r-bands. As established in the preceding section, this requires the hybridization t_{13}.

For $N = 1$, the Hilbert space is small enough for the Hamiltonian to be diagonalized exactly. We do this for a wide range of parameters and calculate the occupation of the orbitals. This is shown for an exemplary set of parameters in Fig. 3.4. We show the occupancy of the lower orbital in the ground state for a total of 3 electrons, two with spin-up and one with spin-down. When there is no hybridization, the system undergoes a series of sharp transitions where electrons abruptly jump from one orbital to another. This occurs whenever the ground state energies of two configuration cross. Beyond $t_{13} \sim 0.3$, these discontinuities disappear; the d-orbital remains occupied over the entire range of Δ. For small Δ, we observe that n_1 is smaller than 1. Because there are 2 r-electrons in total, this means that the outer orbitals start in an inversed occupation, with the supposedly higher orbital being more occupied than the supposedly lower orbital. This unphysical behavior is a result of the interaction. Suppose the upper and lower orbital contain one

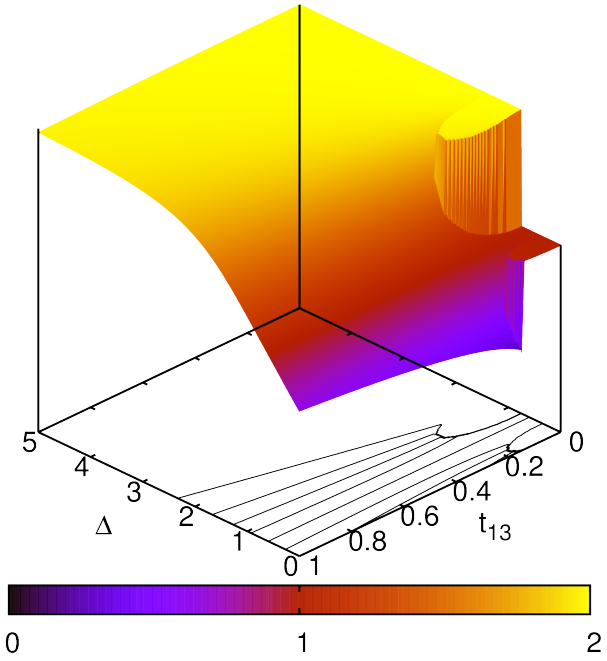


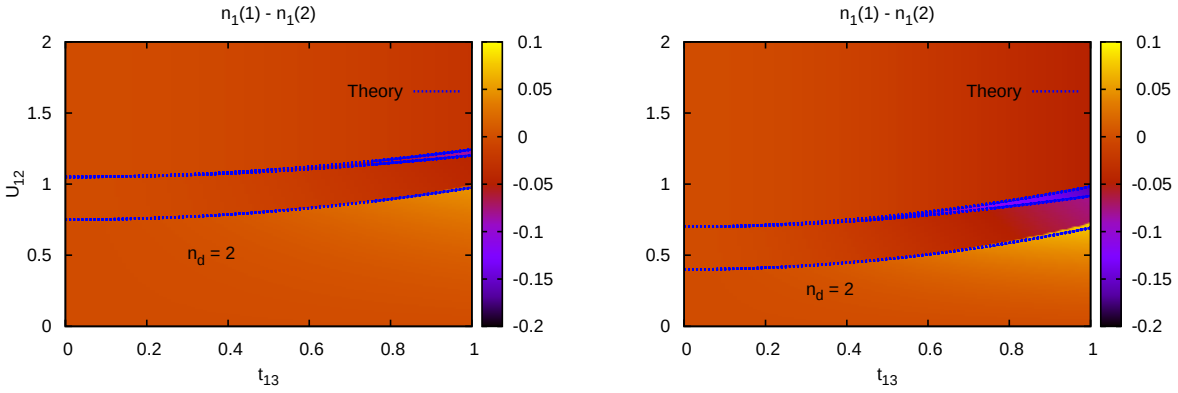
Figure 3.4.: Occupation of the lower orbital for a single site with 3 electrons. The contour lines highlight where the discontinuities occur. Parameters are $U_{11} = 1.5, U_{22} = 0.6, U_{33} = 0.4, U_{12} = 0.7, U_{23} = 0.05$.

electron each. Adding an electron to the lower one costs energy $-\Delta + U_{11} + 2U_{12}$ while adding it to the upper one costs energy $\Delta + U_{33} + 2U_{12}$ (ignoring the hybridization for the moment). Thus, for small delta, orbital 3 is favored over orbital 1 if $U_{33} + 2n_d U_{23} < U_{11} + 2n_d U_{12}$, which is satisfied for all our parameter sets because $U_{33} < U_{11}$ and $U_{23} > U_{12}$. When the band separation Δ is increased, the occupancies approach the limit of independent bands, where the lower and central band are filled while the upper band is empty. The main conclusion is that introducing a hybridization indeed leads to a situation where we have two d-electrons and less than two electrons in the lower band. Hence, whenever a new set of parameters is studied one should first calculate the value of n_2 for $N_e = 2, 3, 4$ and verify that $n_2 = 0, 1, 2$ holds.

Next, we will verify that the d-electrons actively change the occupancy of the outer bands. The parameters mainly responsible for this behavior are U_{12}, U_{23}, and t_{13}. Therefore, we study how n_1, the occupancy of the lower orbital, changes when we move from one to two d-electrons. That is,

$$\delta n_1 = n_1(n_2 = 1) - n_1(n_2 = 2). \tag{3.5}$$

We plot this as a function of both U_{12} and t_{13} in Fig. 3.5 for a parameter set where most other parameters are set to 0, so matters are kept simple. In this case, the resulting on-site Hamiltonian for the r-electrons can be diagonalized exactly (cf. Appendix B). This



(a) Weak case

(b) Strong case

Figure 3.5.: Influence of hybridization t_{13} and intra-orbital interaction U_{12} on electron density. $U_{11} = 0$, $U_{33} = 0$, $U_{23} = 0$, $U_{22} = 0.6$. The weak case has $\Delta = 2.1$, the strong case has $\Delta = 1.4$.

allows us to derive a closed formula for the occupation of the lower orbital. The result is

$$n_1 = 1 + \frac{2\Delta - U_{12}n_2}{\sqrt{(2\Delta - U_{12}n_2)^2 + 4t_{13}^2}}, \tag{3.6}$$

which holds under the condition that the state with n_2 d-electrons is indeed the ground state of the system. Eq. (3.6) allows us to compute δn_1 and compare it to our system. In Appendix B, we also compute the ground state energies of the system in the defect case where d-electrons leave their orbital. By comparing the ground state energies for the different situations, we can obtain formulas for the discontinuities in the occupation of the orbitals (cf. Appendix B.1). These are drawn as dotted lines in Fig. 3.5 and agree well with the data obtained from the simulation.

In the regime where the d-electrons behave as intended ($n_d = 2$ as indicated in Fig. 3.5), we can write down a formula for δn_1 using (3.6). As we cannot simplify the resulting expression, we apply a Taylor expansion for small U_{12}/Δ and t_{13}/Δ. The result is

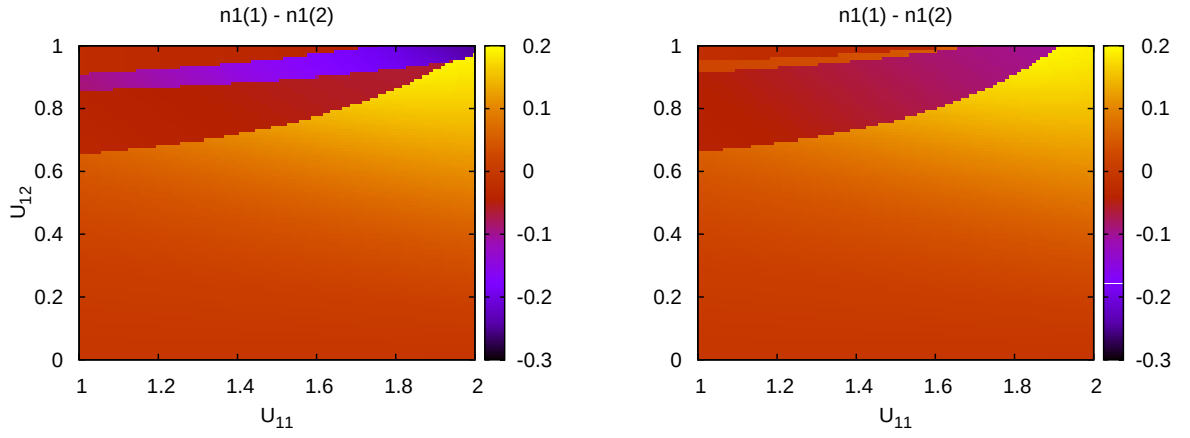
$$\frac{\delta n_1}{\Delta} \approx \left(\frac{t_{13}}{\Delta}\right)^2 \left(\frac{U_{12}}{2\Delta} + \frac{9}{8}\left(\frac{U_{12}}{\Delta}\right)^2\right). \tag{3.7}$$

From this, we learn that first, the influence of other bands is generally stronger when the energy separation Δ is smaller, which comes as no surprise. Second, we learn that U_{12} is stronger in driving the screening, as it enters (3.7) with both a linear and a quadratic term, whereas t_{13} only occurs in second order. The role of t_{13} is to determine how strong the screening can become before the interaction leads to electrons leaving the d-band, as indicated by the discontinuities discussed above. We conclude that the model behaves as intended: There exist regimes for the parameters where electrons can be added to the central band so that the r-electrons rearrange. The strength of this effect depends on the hybridization and, more strongly, on the intra-band density-density interaction.

In the calculation above, the other interactions were set to 0. We will now discuss their role. As in this case an analytical diagonalization of the Hamiltonian becomes difficult, we are restricted to a qualitative discussion. U_{23} is the counterpart to U_{12}. This is seen by rewriting H_{rd} for a single site by using that the total number of r-electrons is fixed at 2 per site, $n_1 + n_3 = 2$.

$$H_{rd} = U_{12}n_1n_2 + U_{23}n_2n_3 = (U_{12} - U_{23})n_1n_2 + 2U_{23}n_2. \tag{3.8}$$

Thus, U_{23} renormalizes both U_{12} and the orbital energy of the d-electrons. The on-site energies U_{11} and U_{33} shift the energy for double-occupancy of the outer orbitals. U_{11} affects the energy cost of moving an electron from the initially full lower orbital to the initially empty upper orbital. The higher U_{11}, the more energy is gained by moving the electron, which enhances the effect of both t_{13} and U_{12}. This is diminished by U_{33}, which works against doubly occupying orbital 3. The role of U_{11} and U_{33} is shown in Fig. 3.6. First, we note that discontinuities are, again, due to d-electrons leaving their



(a) $U_{33} = 0$ (b) $U_{33} = 0.2$

Figure 3.6.: δn_1 as a function of U_{11} and U_{12} for different values of U_{33}. $U_{22} = 0.6, U_{23} = 0, t_{13} = 0.5, \Delta = 1.6$. (a) $U_{33} = 0$, (b) $U_{33} = 0.2$.

band. Second, we observe that the shape of the first discontinuity indicates that increasing U_{11} serves to stabilize the situation. This is because U_{11} leads to a reduced occupation of the lower orbital, thereby reducing the influence of U_{12}. While the direct effect of U_{11} on the screening behavior is small compared to that of U_{12}, it leads to larger screening when combined with higher values of U_{12}. For small values of U_{11}, the interaction U_{33} has negligible effect. This is expected, as in that regime the occupation of the upper orbital is low. For larger values of U_{11}, however, U_{33} has the effect of further stabilizing the system, as can be seen from the boundary of the discontinuity. Without U_{33}, the d-electrons would jump into the upper orbital. This is made energetically unfavorable by increasing U_{33}.

3.3.2. Single-band Properties

Since our motivation is to reduce the full model to an effective single-band model, we now discuss some general properties of the one-dimensional single-band Hubbard model.

Non-interacting Case

Let us begin with the non-interacting case, $U = 0$. There, the electrons are independent of each other, except for the Pauli exclusion principle. The eigenenergies are obtained by reversing (1.8), which yields

$$\varepsilon_{\mathbf{k}} = -2t\cos(ka), \tag{3.9}$$

where a is the lattice constant and k is restricted to the first Brillouin zone, i.e., $k \in (-\frac{\pi}{a}, \frac{\pi}{a}]$. The type of boundary conditions employed further restricts the possible values of k. The two most common cases are periodic and antiperiodic boundary conditions. For periodic boundary conditions we have

$$k = \frac{2\pi}{Na} n, \lfloor \frac{N}{2} \rfloor, \tag{3.10}$$

while for antiperiodic ones we have

$$k = \frac{2\pi}{Na}(n - 1/2). \tag{3.11}$$

When k is restricted to the first Brillouin zone, we have

$$n = -\lfloor \frac{N-1}{2} \rfloor \ldots \lfloor \frac{N}{2} \rfloor. \tag{3.12}$$

While we expect the precise choice of boundary conditions to be irrelevant in the thermodynamic limit of infinitely large systems, they play a crucial role in small systems. One important aspect is the degeneracy of the many-body ground state. In the non-interacting state, we just have to fill the energy levels in ascending order with two electrons each. However, since the cosine is an even function, we have a twofold degeneracy for all k except $k = 0$ and $k = \pi/a$. This is illustrated in Fig. 3.7. Each energy level can accommodate up to two electrons. Thus, for periodic boundary conditions (pbc) the ground state is degenerate, while it is unique for antiperiodic boundary conditions (abc). The general rule is that for an even number of sites, a non-degenerate ground state is obtained with antiperiodic boundary conditions if the number is a multiple of four, and with periodic boundary conditions otherwise. The reason is as follows. For periodic boundary conditions, there is an odd number of (partially) occupied orbitals. The ground state is degenerate when these orbitals are not completely filled. Thus, an odd number of electrons of each spin type is required to achieve a non-degenerate state. At half filling, this situation is obtained for an even number of sites that is not a multiple of four. For antiperiodic boundary conditions the argument is analogous. For an odd number of sites, the ground state at half filling will always be degenerate, because one of $N_\uparrow$ and $N_\downarrow$ will be even while the other will be odd. Thus, for either type of boundary conditions degeneracy occurs.

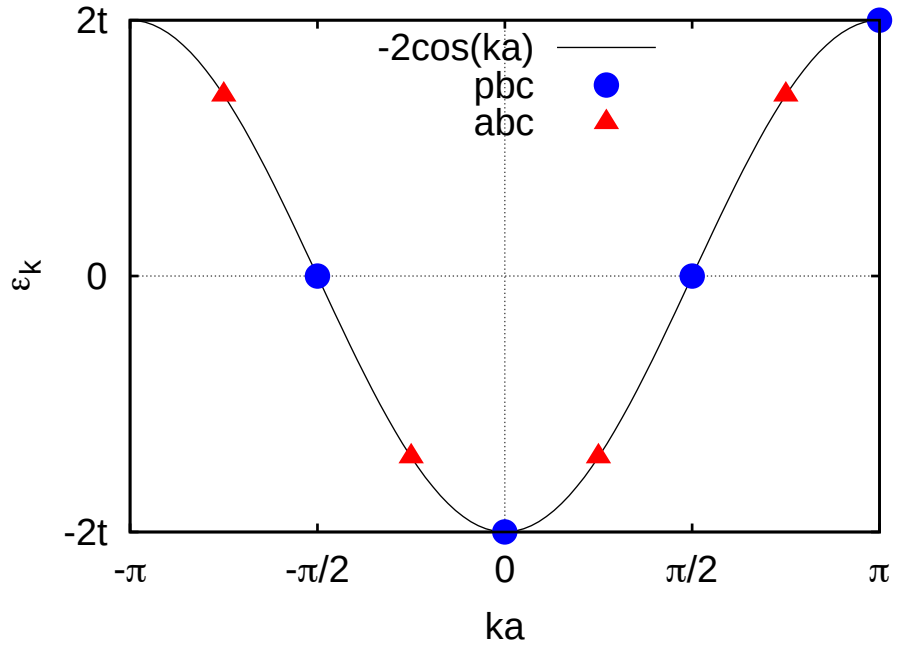


Figure 3.7.: Energy band (solid line) with the discrete k-values obtained for periodic (pbc) and antiperiodic (abc) boundary conditions for a system with $N = 4$ sites.

Interacting Case

Let us now move to the case with nonzero Hubbard U. There are two limiting cases, expressed by the ratio U/t. For $U/t \to \infty$, we speak of the *atomic limit*. Because of the strong repulsive interaction, the system abhors doubly occupied sites. For half filling, every lattice site will contain exactly one electron. As moving one electron to a neighboring site costs energy U, electron transport is strongly inhibited. This is the so-called Mott insulating state [47]. In their paper, Mott and Peierls explained why some transition metal oxides were insulating despite their band structure suggesting they should be metallic. By treating the hopping as a perturbation, one can show that this state is antiferromagnetic. This is a result of virtual hopping; if the spins are aligned antiferromagnetically, they can hop at least virtually, thereby lowering the kinetic energy. For parallel spins, the Pauli exclusion principle prevents such virtual hopping.

The other limit is the *band limit*, characterized by $\frac{U}{t} \to 0$. This is the non-interacting case, which was discussed in the previous section. The resulting system is metallic for $U = 0$. In the one-dimensional case, it was shown by Lieb and Wu [12] that the Mott insulating state is obtained for every $U > 0$. For higher dimensions, the system remains metallic for small U and exhibits a Mott transition from metal to insulator at some finite U. This transition is clearly visible in Fig. 3.8. For $U = 0$, we see 6 distinct peaks that correspond to the 6 twofold degenerate eigenstates corresponding to the 12 k-points. For small U, the shape of the spectral function is still close to the non-interacting limit, but increasing U opens a Mott gap. The spectra clearly show the formation of a lower and

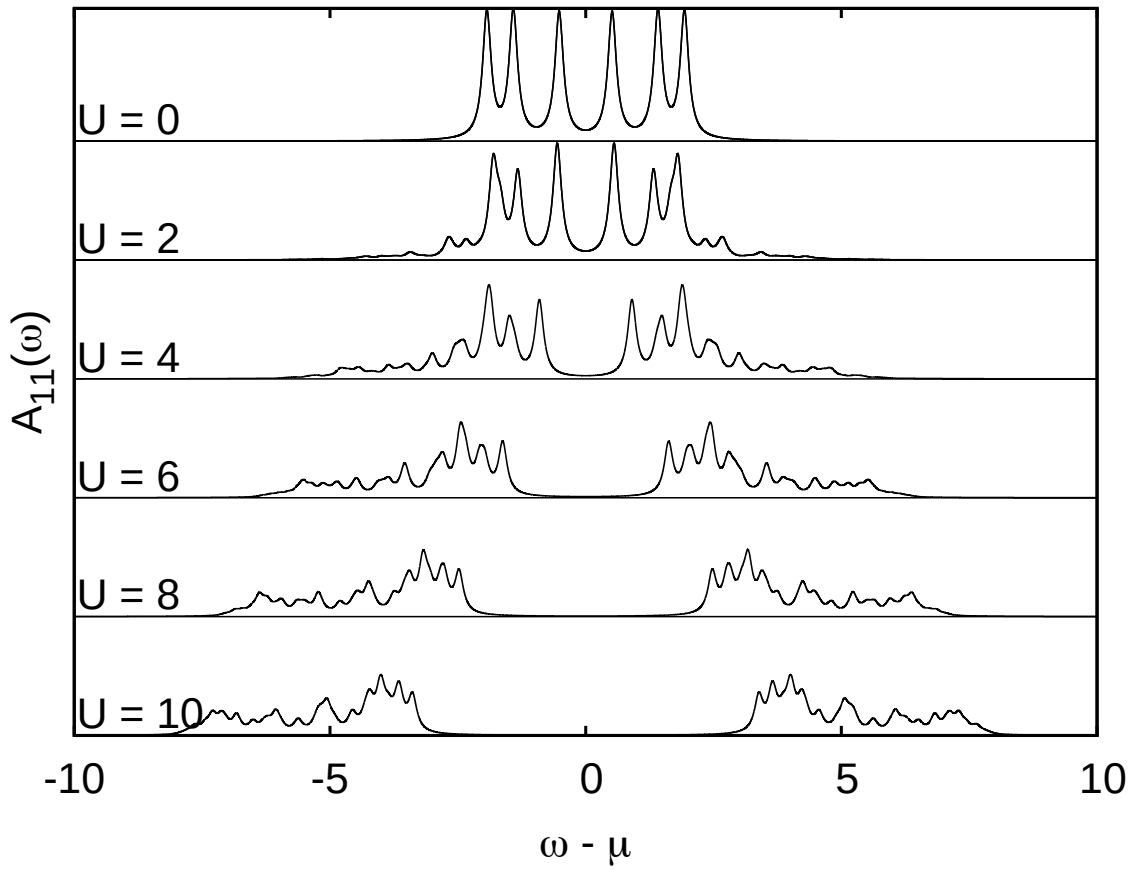


Figure 3.8.: Spectral functions of a one-dimensional Hubbard chain with $N = 12$ sites at half-filling and antiperiodic boundary conditions. t is -1, U is as shown.

an upper Hubbard band, whose centers are approximately U apart.

3.3.3. Multi-band Properties

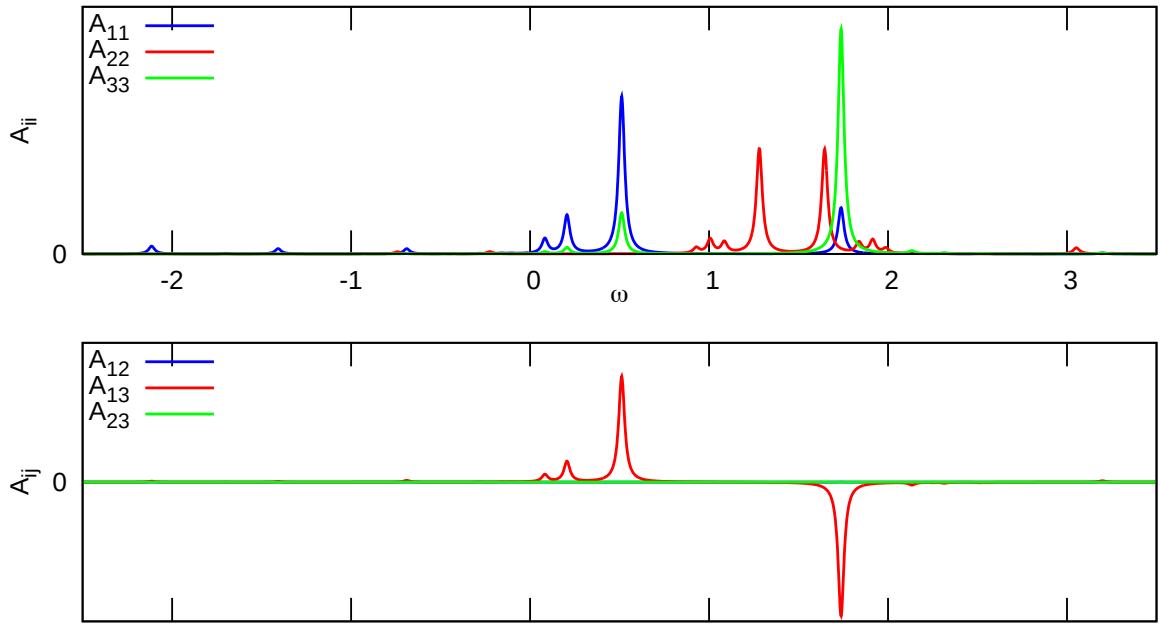
In Sec. 3.3.1 we already investigated the properties of the full system for one site. We conclude this chapter with a discussion of the spectral function of the full system for more than one site.

First, we look at the size of systems we can handle computationally. Let N be the number of orbitals and L the number of lattice sites. As we have three orbitals per site, we have $N = 3L$. For symmetry among the up- and down-electrons, we want L to be even, which further restricts N to multiples of 6. For $L = 2$, solutions can be obtained using direct diagonalization. For $L = 4$, they can be computed comfortably on a modern laptop or desktop computer. The next step is $L = 6$, which already requires MPI parallelization, as a many-body vector for such a system requires 17 GB of memory (cf. Tab. 2.1). Systems with $L = 8$ sites are currently out of reach. They require 54

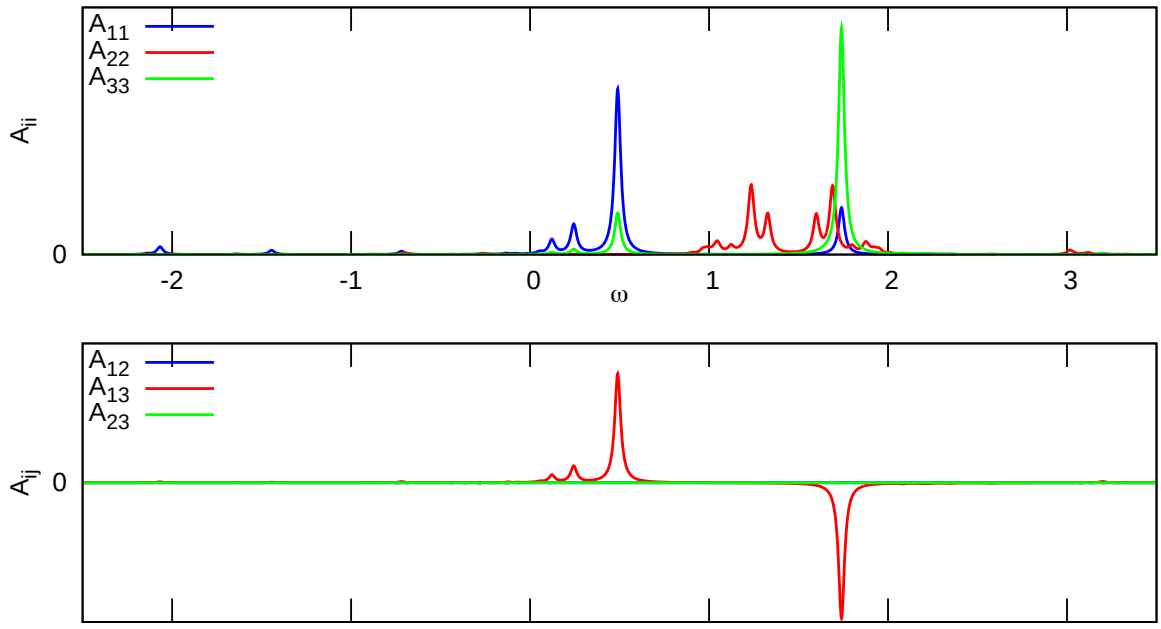
TB of memory for each many-body vector, while JUGENE has a total memory of 144 TB available. We expect systems of that size to be manageable after the next major upgrade of the supercomputing facilities at Juelich.

Second, we mention a numerical issue that was observed in the first test runs. As H not only preserves the total number of r-electrons but also the number of r-electrons per site, we are left with whatever was the expectation value for that number in the (random) starting vector. For a sufficiently large system, this expectation value will be close to the correct value of 2 r-electrons per site, as this is the average over all possible configurations. However, the observed deviation for a system with $L = 4$ sites, for example, is still larger than the numerical accuracy of the ground state energy. Moreover, the number of d-electrons will also not come out exactly as desired but depend on the initial starting vector. This leads to several undesirable properties of the ground state vector obtained from the Lanczos algorithm. First, translational invariance is lost as the number of r-electrons per site differs. Second, reproducibility is lost because the distribution of electrons among r- and d-orbitals depends on the randomized starting vector. Fortunately, the problem can easily be solved by adding a very small hopping t_{23} or t_{12} to the system. In most of our calculations, we chose $t_{23} = t_{22}/1000$. This allows the Hamiltonian to freely move electrons between all three bands without significantly changing the physics of the system. When the parameters are chosen correctly, this results in states with the desired number of d- and r-electrons and with translational invariance.

Finally, we present and discuss the full spectral function $A_{ij}(\omega)$ of a three-band system with $L = 4$ and $L - 6$ sites (cf. Fig. 3.9). The diagonal and off-diagonal parts of the spectral function are shown in the upper and lower parts of the plot, respectively. We observe that the d-electrons have a spectral function, A_{22}, that looks similar to the ones of a single-band Hubbard chain. For the parameters used, the spectral functions of the r-electrons have significant overlap with the spectral function of the d-electrons. Therefore, we expect renormalization effects to occur. Note that the Mott gap in A_{22} is not centered at $U_{22}/2$. This is because the influence of the r-electrons changes the effective on-site energy and Hubbard U of the d-electrons. We will discuss this in much detail in the next chapters. Peaks in the off-diagonal spectral functions indicate inter-band excitations. We see that only A_{13} exhibits such peaks, as the Hamiltonian has no matrix elements connecting the d- and r-orbitals. Finally, we note that A_{22} is more defined when 6 sites are used. This is because for more sites, the resolution in k-space improves. Also, for antiperiodic boundary conditions, as used for $L = 4$, the peak at $k = 0$ is always missing, so the Mott gap does not show as clearly. Because the overall structure is, however, very similar for $L = 4$ and $L = 6$ sites, we will use 4 sites for most of our calculations to reduce costs and save time.



(a) $L = 4sites$



(b) $L = 6sites$

Figure 3.9.: Full spectral function A_{ij} of a three-band Hubbard chain with (a) $L = 4$ sites, antiperiodic boundary conditions, and (b) $L = 6$ sites, periodic boundary conditions. Parameters for both are $U_{11} = 1.5$, $U_{22} = 0.6$, $U_{33} = 0.4$, $U_{12} = 0.7$, $U_{23} = 0.05$, $t_{22} = -0.1$, $t_{13} = -0.5$, $\Delta = 1.4$

4. Screening

We will now begin the discussion of screening and renormalization effects in our Hubbard model. In Sec. 4.1, a simple approach to calculating a renormalized Hubbard U is given. A systematic derivation follows in Sec. 4.2.

To visualize the renormalization of the Hubbard U due to the r-electrons, we compare the Green's Function of the full model with the Green's Function of a single band model with $U = U_{22}$. The result is shown in Fig. 4.1. For this and future example calculations, we use the parameters given in Tab. 4.1. In addition to the small extra peaks at very high and very low energies, the d-band in the full system also has a Mott gap that is significantly smaller than that of a single band system with the same Hubbard U. Apart from that, both spectral functions are similar in shape. This indicates that it might be possible to reproduce the full spectral function by using a Hubbard U smaller than U_{22} in the effective single band model.

Table 4.1.: Parameter set for exemplary calculations.

U_{11}	1.5	U_{12}	0.7	t_{22}	-0.1
U_{22}	0.6	U_{23}	0.05	t_{13}	-0.5
U_{33}	0.4	Δ	1.4		

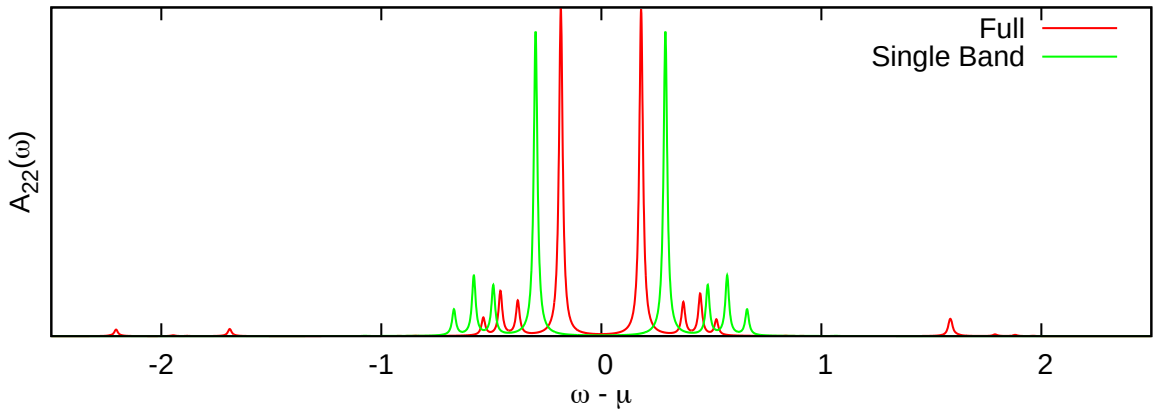


Figure 4.1.: Spectral function of the central band of the full system compared to the spectral function of an effective single band system with $t = t_{22}$ and $U = U_{22}$. Parameters are taken from Tab. 4.1.

4.1. Static Screening

We now present a very simple way of calculating an effective Hubbard U, which we will call *static screening*. Let $\mathcal{E}_i$, $i = 0, 1, 2$, be the ground state energy for a single site of the full system in the atomic limit with $0, 1$ or 2 electrons in the central orbital, respectively. These energies can be obtained via exact diagonalization of the many-body Hamiltonian for a single site. Under the assumption that the r-electrons will always be in the ground state with respect to the number of central band electrons, the energy for adding a d-electron to a previously empty orbital is $\mathcal{E}_1 - \mathcal{E}_0$ and the energy for adding another d-electron to that site is $\mathcal{E}_2 - \mathcal{E}_1$. Now, we compare this to the Hubbard model. In the single band Hubbard model, adding one electron to a previously empty sites costs the orbital energy ε of that site, and adding an additional electron costs the orbital energy plus the Coulomb interaction energy U, as shown in Fig. 4.2. We therefore identify

$$\begin{aligned} \varepsilon &= \mathcal{E}_1 - \mathcal{E}_0 \\ U &= \mathcal{E}_2 - 2\mathcal{E}_1 + \mathcal{E}_0. \end{aligned} \tag{4.1}$$

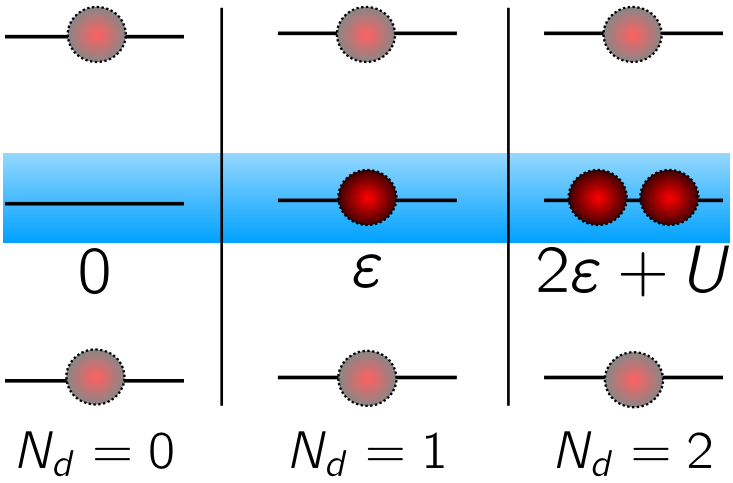


Figure 4.2.: Energies associated with adding d-electrons to a site, and their relation to the Hubbard parameters ε and U.

For a system with fixed number of electrons N_{el} the orbital energy ε leads to a constant shift of the total energy of size $N_{\mathrm{el}}\varepsilon$ and could be ignored while computing the ground state. However, it is necessary to retain the correct total energy when comparing the single band model to the full model, as it affects the system's chemical potential.

4.1.1. Results

We apply this method of calculating effective parameters ε and U to the system from Fig. 4.1 and obtain $\varepsilon = 1.23$, $U = 0.457$. This effective U is reduced from $U_{22} = 0.6$ by about 24%. The resulting spectral function is shown in Fig. 4.3. With the renormalized

U, the estimated Mott gap is now very close to that of the full system. We can state that static screening is capable of providing a simple way of calculating an effective Hubbard U. For the parameters of Tab. 4.1, we plot the resulting difference between the screened

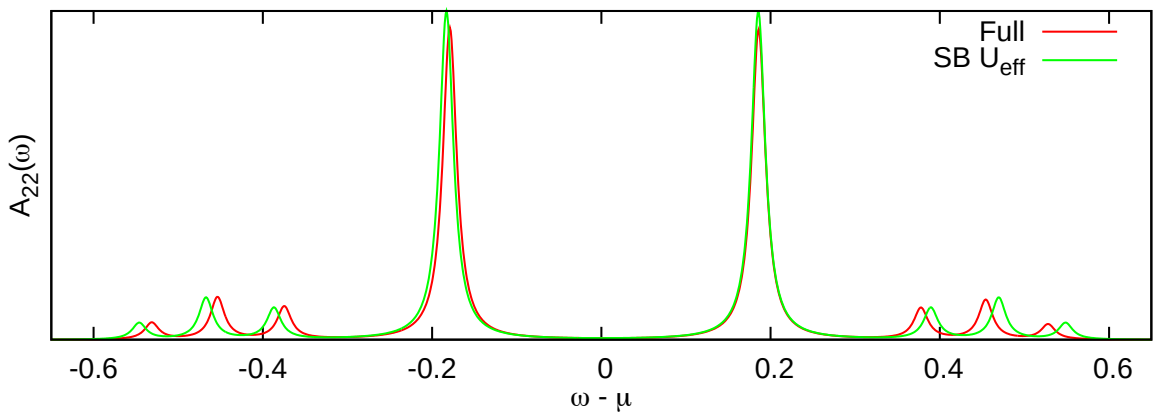


Figure 4.3.: Spectral function of an effective single band Hubbard model with U obtained via static screening. Parameters are the same as in Fig. 4.1, η is 0.01 in this and any future plots unless stated otherwise.

U from the bare U_{22} in Fig. 4.4. In the limit of large Δ, U barely changes, since in this

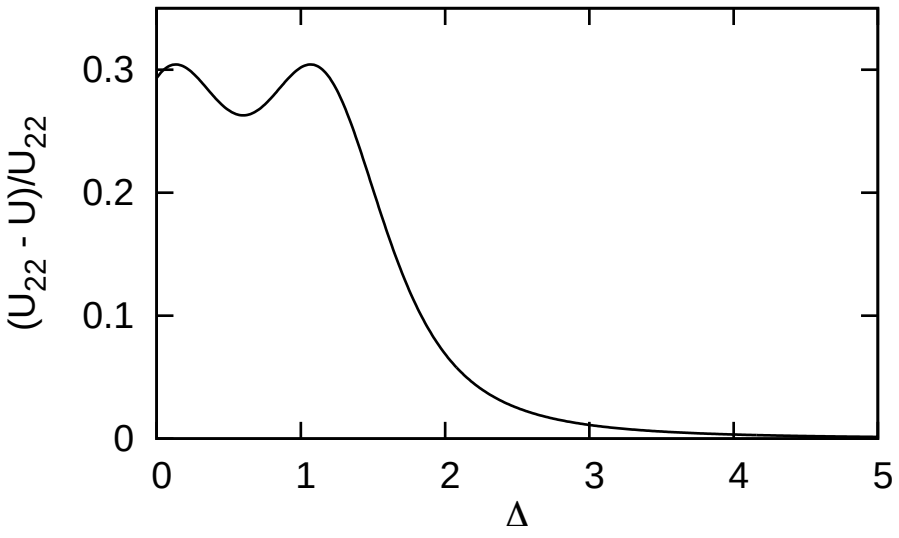


Figure 4.4.: Relative difference between the screened U from the bare U_{22} as a function of the band separation Δ. Other parameters are from Tab. 4.1.

case the bands are energetically well separated so that transitions of r-electrons from the lower to the upper band cost too much energy. For intermediate Δ, screening becomes

more and more efficient and approaches values of up to 30%. In the regime of small Δ, unphysical behavior can occur. As discussed in Sec. 3.3.1, the occupation of the r-orbitals can become inverted in this case. The oscillation below $\Delta = 1$ is due to this orbital inversion.

4.1.2. Chemical Potential

When we compare the spectral function of the d-electrons in the full system with that of electrons in a single-band system, we are comparing systems with two different Hilbert spaces. Therefore, we have to compare the centered spectra, $A(\omega - \mu)$, rather than just the $A(\omega)$. To this end, we have to calculate the chemical potential, μ.

To obtain the correct total energy and chemical potential of the d-electrons, we must include the energy of the r-electrons. For a single site, we can write the wavefunction of the r-electrons as

$$|\psi_r\rangle = e\,|0\rangle + p\,|1\rangle + d\,|2\rangle \tag{4.2}$$

where $|n\rangle$ denotes the r-electron ground state obtained when there are n d-electrons on the site (cf. Fig. 4.2). The expectation value of $H_r + H_{rd}$ is then, by definition,

$$\mathcal{E}_r = \langle\psi_r|\, H_r + H_{rd}\,|\psi_r\rangle = E\mathcal{E}_0 + P\mathcal{E}_1 + D(\mathcal{E}_2 - U_{\text{bare}}) \tag{4.3}$$

where $E = ee^*$, $P = pp^*$, $D = dd^*$, i.e., E, P and D are the probabilities for a central orbital containing zero, one, or two electrons. The probabilities can be obtained either from the ground state of the full system or within the Gutzwiller Approximation (GWA) [11, 48, 49]. We can rewrite (4.3) by replacing $\mathcal{E}_1$ and $\mathcal{E}_2$ with $\varepsilon + \mathcal{E}_0$ and $U + 2\varepsilon + \mathcal{E}_0$, respectively. This yields

$$\mathcal{E}_r = \mathcal{E}_0 + \varepsilon n_d - D(U_{\text{bare}} - U). \tag{4.4}$$

Using these energies leads to a good approximation for the total energy and a chemical potential that is close to that of the full system.

For the special case of the one-dimensional single-band Hubbard model at half filling, the chemical potential has the simple form

$$\mu = \varepsilon + \frac{U}{2}. \tag{4.5}$$

When the single-band description of the d-electrons is appropriate, (4.5) already is a very good approximation to the chemical potential of the full system.

4.1.3. Overview of the Parameter Regimes

Depending on the parameters, the system can exhibit weak screening, strong screening or too strong screening that renders the single-band description of the d-electrons unphysical. We will now explore the parameter regimes that lead to these kinds of behavior. First, we introduce a quantitative measure for the agreement between spectral functions and then use it to identify regimes where the single band approximation is valid and regimes where it breaks down.

Error Measure for Spectral Functions

Let $||\cdot||_1$ denote the $\mathcal{L}^1$-norm of a function, that is,

$$||f(\omega)||_1 = \int_{-\infty}^{\infty} |f(\omega)|. \tag{4.6}$$

As a measure for the discrepancy between two (diagonal) spectral functions A_1 and A_2, we define

$$E_\eta[A_1, A_2] = \frac{1}{2}||A_1(\omega) - A_2(\omega)||_1, \tag{4.7}$$

where η is the Lorentz broadening factor used in the spectral functions. We choose the $\mathcal{L}^1$ norm, because for diagonal spectral functions, A_{ii}, there is the sum rule

$$||A(\omega)||_1 = \int_{-\infty}^{\infty} A_{ii}(\omega) = 1. \tag{4.8}$$

This allows us to easily normalize the error measure. Clearly, it is largest for spectral functions that have disjunct support, i.e., for any given ω only one of A_1 or A_2 is nonzero. In that limiting case, we can write the norm as

$$E[A_1, A_2] = ||A_1(\omega)||_1 + ||A_2(\omega)||_1 = 2.$$

This limiting case is obtained for $\eta \to 0$, where the spectral functions become a collection of δ-peaks and the slightest shift leads to misalignment. With the prefactor of $1/2$, we then have $0 \leq E_\eta \leq 1$.

As the convergence factor η determines the peak width of the spectral functions, it also influences the error measure. We show this in Fig. 4.5 for our example system. For large η, we obtain broad peaks. Consequently, minor differences in A_1 and A_2 are smeared out and the resulting error measure is small. For small η, in contrast, peaks are very

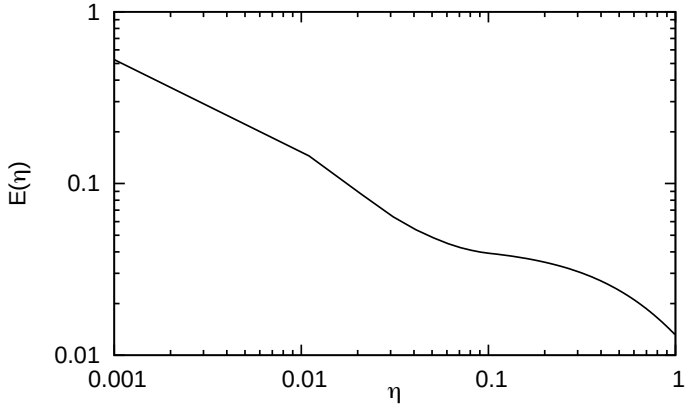


Figure 4.5.: Error measure for the full and the effective spectral function obtained from static screening. Parameters are from Tab. 4.1, corresponding to the spectral function shown in Fig. 4.3.

narrow so that even small misalignments between A_1 and A_2 result in large errors. More importantly, small improvements in the alignment are not reflected in the error measure. Hence, when selecting a value for η we aim for a peak width that is slightly smaller than the smallest peak separation. In most situations, we found that $\eta = 0.01$ is a good choice.

Regimes

The interaction parameter regimes we will investigate in this section are shown in Tab. 4.2. In the future, we will refer to these parameter sets by the abbreviations given in the table.

(a) Strong interactions (Sr)

$U_{11} = 1.5$	$U_{12} = 0.7$
$U_{22} = 0.6$	$U_{23} = 0.05$
$U_{33} = 0.4$	$t_{13} = -0.5$

(b) Weak r-interactions, medium coupling, strong d-interaction (WrMcSd)

$U_{11} = 0.15$	$U_{12} = 0.4$
$U_{22} = 0.6$	$U_{23} = 0.05$
$U_{33} = 0.04$	$t_{13} = -0.5$

(c) Weak r-interactions and coupling, strong d-interaction (WrWcSd)

$U_{11} = 0.15$	$U_{12} = 0.07$
$U_{22} = 0.6$	$U_{23} = 0.00$
$U_{33} = 0.04$	$t_{13} = -0.5$

(d) Weak r-interactions, weak coupling, weak d-interactions (WrWcWd)

$U_{11} = 0.15$	$U_{12} = 0.07$
$U_{22} = 0.1$	$U_{23} = 0.00$
$U_{33} = 0.04$	$t_{13} = -0.5$

Table 4.2.: Parameter sets with different interaction strengths. $\mathbf{t_{22}}$ **is** -0.1 **for all calculations**. The abbreviations are used to reference these sets.

The resulting renormalization of U_{22} for these four cases is illustrated in Fig. 4.6. From this plot, we can already identify several regimes of interest. For large Δ, the screening is smooth and approximately follows a power law, indicated by the linear slopes in the logarithmic plot. The slope indicates a decay that goes approximately like Δ^{-3}. This can be explained by turning again to the exact diagonalization of the on-site Hamiltonian for the special case where only U_{12} and U_{22} are non-zero (cf. Appendix B.3).

The kinks in the curves for parameter sets (WrMcSd) and (WrWcSd) are due to d-electrons leaving their orbital (cf. Sec.3.3.1, where we discussed this behavior). In the region where this has happened, screening appears to be quite strong. This is because the on-site energy $\mathcal{E}_2$ no longer contains the contribution U_{22} from the doubly occupied d-site and therefore underestimates the energy necessary to create a such a doubly occupied site. Consequently, $U_{22} - U$ is overestimated. One should bear in mind, however, that the U obtained in this situation will yield poor agreement with the full spectral function. Finally, for some Δ the screening strength reaches a plateau after which it does not change significantly. In this regime, we expect the single band picture not to be valid any more, as the energy separation Δ becomes smaller than the band width of the central band. It will be worthwhile to study just when this breakdown occurs.

We show the computed error measure with $\eta = 0.001$ in Fig. 4.7. We use such a relatively low value for η, because otherwise numerical problems may arise with to the numerical integration of $|A_1(\omega) - A_2(\omega)|$ that arise when the error measure becomes very

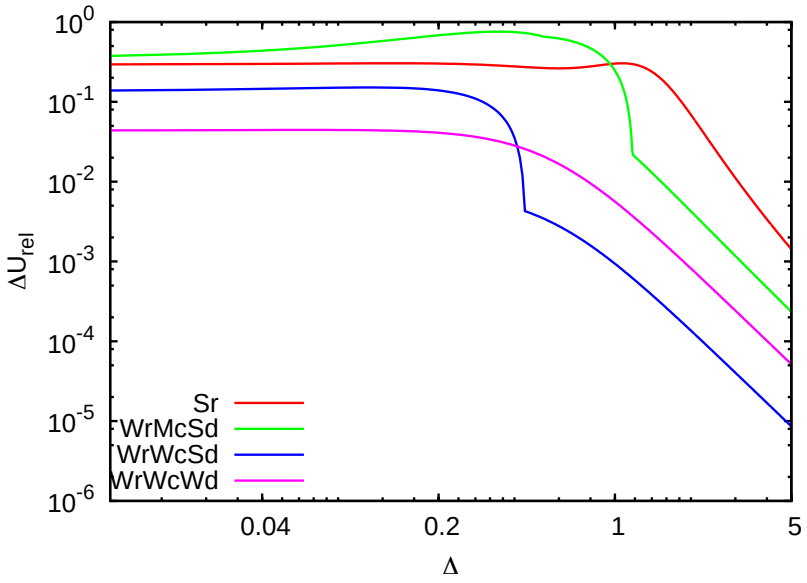


Figure 4.6.: Relative renormalization of U_{22} due to the outer bands. The letters refer to those in Tab. 4.2.

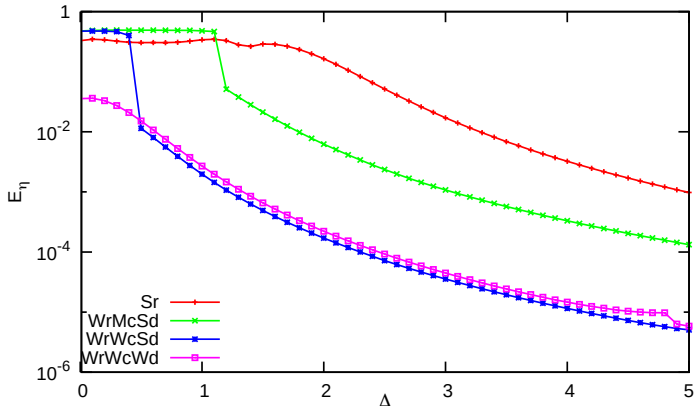


Figure 4.7.: Error measure E_η with $\eta = 0.001$ for the four parameter sets from Tab. 4.2. Such a small value for η is used because for weak coupling and large Δ, where the error measure is on the order of 10^-5 and below, differences in the spectral function are lost in the numerical error if the peaks are too broad.

small. When E_η is of the order 10^{-5}, even small misalignments due to, e.g., a wrong chemical potential or other numerical errors,

The behavior of E_η is consistent with the observations from Fig. 4.6. We observe that the discontinuity in the screened U coincides with a sudden and large increase in the discrepancy between the full spectral function and the effective single band spectral function. This confirms that the U obtained from this unphysical case leads to poor agreement; the error measure is almost at its maximum of 1. In the regime of large Δ, the agreement between full and single band spectra becomes, overall, better, with weaker

interactions and weaker coupling leading to lower values of E_η. This is expected, as the limit $\Delta \to \infty$ corresponds to independent orbitals.

The error measure gives us a general idea of when the effective single band model is an appropriate description of the d-electrons. We conclude this section by a discussion of four parameter sets taken from different regimes of Fig. 4.6. The first three plots are all for parameter set (Sr) from Tab. 4.2 but different values of Δ.

- We begin with large Δ, where the effective single band model describes the spectrum of the d-electrons perfectly, with no visible deviation.

- The next plot is for $\Delta = 1.4$ and was shown before (cf. Fig. 4.3). The Mott gap is reproduced correctly, but the minor peaks are misaligned. This indicates that we overestimate the hopping parameter t, which here, as in all calculations in this section, was taken to be the bare value, t_{22}. A thorough examination of this effect is given in the next section.

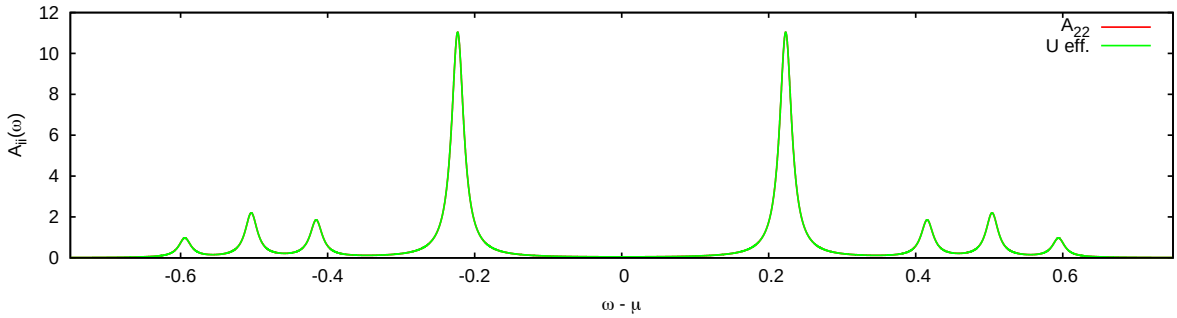
- The third plot is for a very small Δ of 0.2. Here, the supposedly lower band lies above the supposedly upper band, but this orbital inversion (cf. Sec. 3.3.1) appears not to affect the validity of our single band description of the d-electrons. This may surprise at first, but note that in our calculation of the Hubbard U no explicit use is made of the structure of the r-bands. We plot the other two spectral functions, A_{11} and A_{33} in Fig. 4.9(a). We see that, despite Δ being small, the spectral functions of the r-electrons have similar overlap with that of the d-electrons as found for $\Delta = 1.4$ (Fig. 3.9). As the static screening approach to determine U works well in this case, it is no surprise that it also works well for $\Delta = 0.2$.

- Finally, the fourth plot shows the spectral functions for the case where a d-electron leaves its orbital. In this case the procedure to determine U is not justified anymore, so we find, as expected, that the agreement between A_{22} of the full system and that of the effective single band system is poor. First, the reduction of U due to screening is overestimated. Second, the chemical potential of the full system's d-electrons cannot be approximated by $\varepsilon + U/2$ with ε and U taken from the static approach. When we take a look at the other spectra, A_{11} and A_{33}, we find in the gap of A_{22} a pronounced peak belonging to the r-electrons (cf. Fig. 4.9(b)). This indicates that the single band picture really is inappropriate for describing the situation.

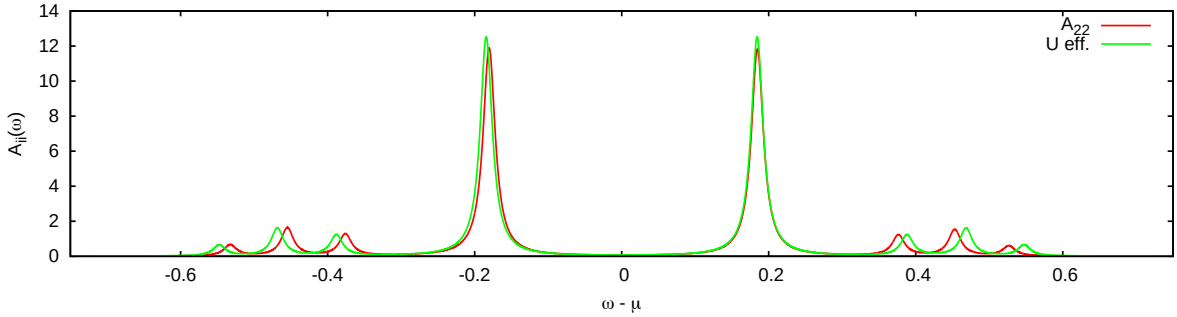
It is important to note that the gaps in the spectra shown in Fig. 4.8 are not the true Mott gaps of the systems but rather include a part that is due to finite size effects. The plots have been done for $L = 4$ sites and antiperiodic boundary conditions. When these are used, there are no peaks at $\omega - \mu = 0$. The finite size gap for a system with 4 sites can be calculated to

$$2t|\cos(\frac{1}{4}\pi) - \cos(\frac{3}{4}\pi)| \approx 0.28t. \tag{4.9}$$

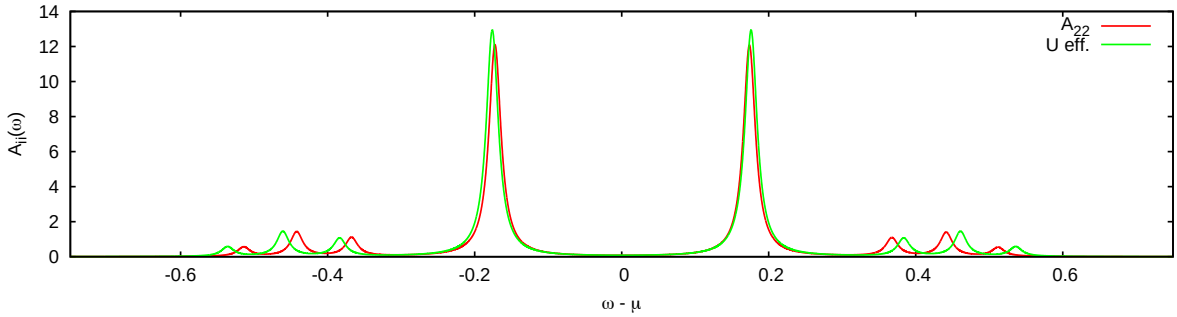
When the Mott gap is sufficiently large, it can be obtained from the spectral function. Otherwise, it is hidden in the finite size gap. Therefore, we also show calculations for a larger system of $L = 6$ sites in Fig. 4.10. We see that the large gap in the previous



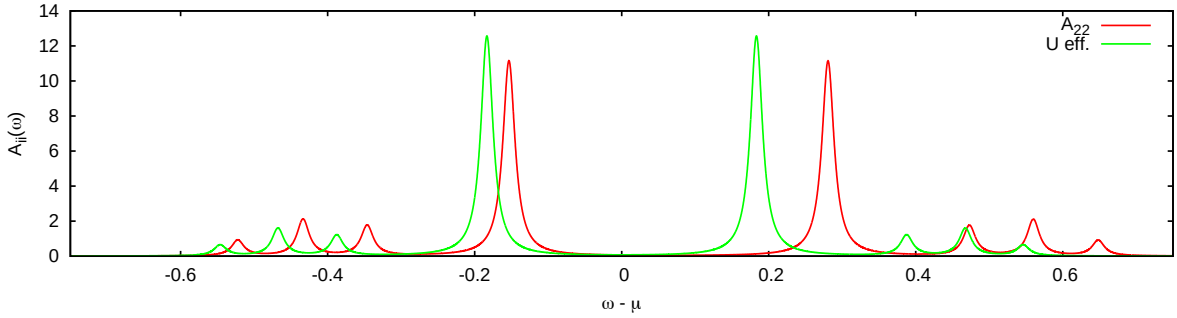
(a) Sr, $\Delta = 5$



(b) Sr, $\Delta = 1.4$

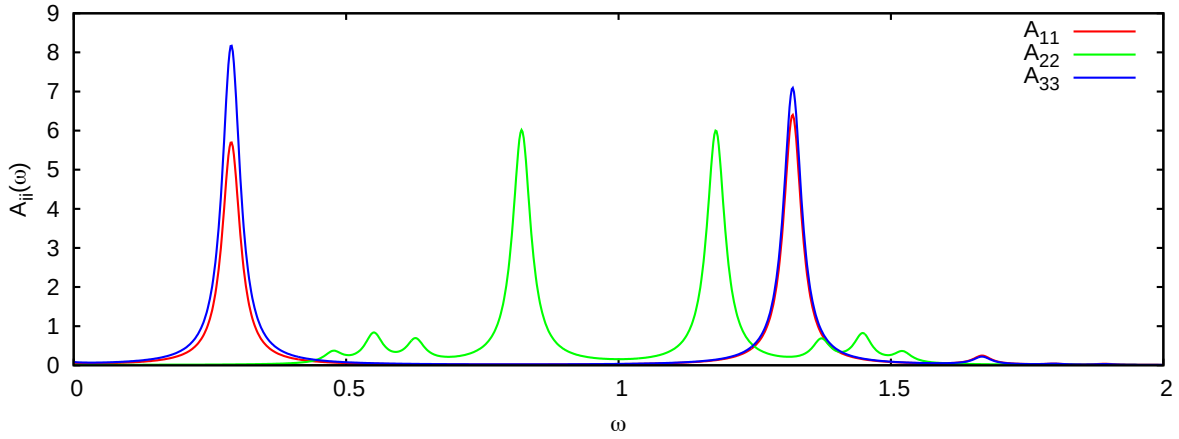


(c) Sr, $\Delta = 0.2$

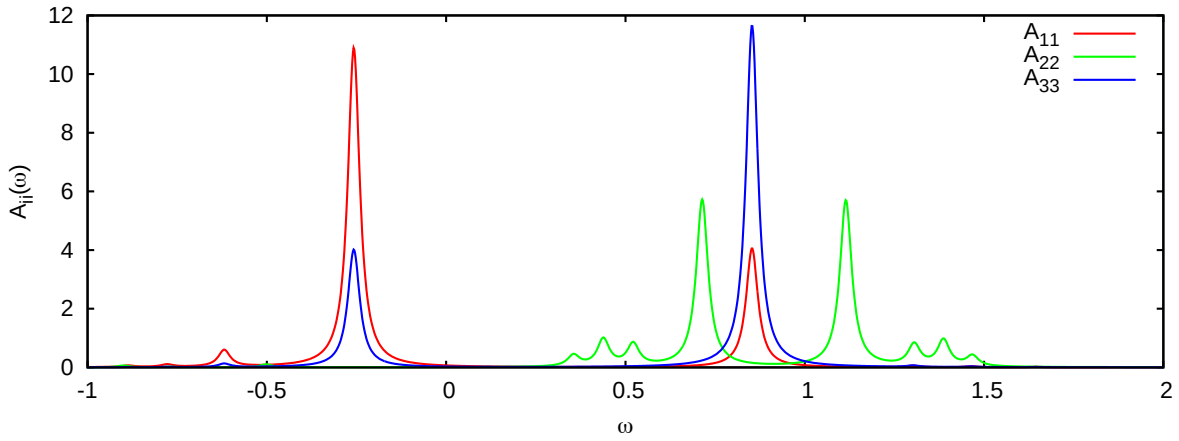


(d) WrMcSd, $\Delta = 1$

Figure 4.8.: Spectral functions for different cases.



(a) $\Delta = 0.2$, Params (Sr)



(b) $\Delta = 1.0$, Params (WrMcSd)

Figure 4.9.: Full diagonal spectral functions for (a) $\Delta = 0.2$ and parameter set (Sr) and for (b) $\Delta = 1.0$, parameter set (WrMcSd).

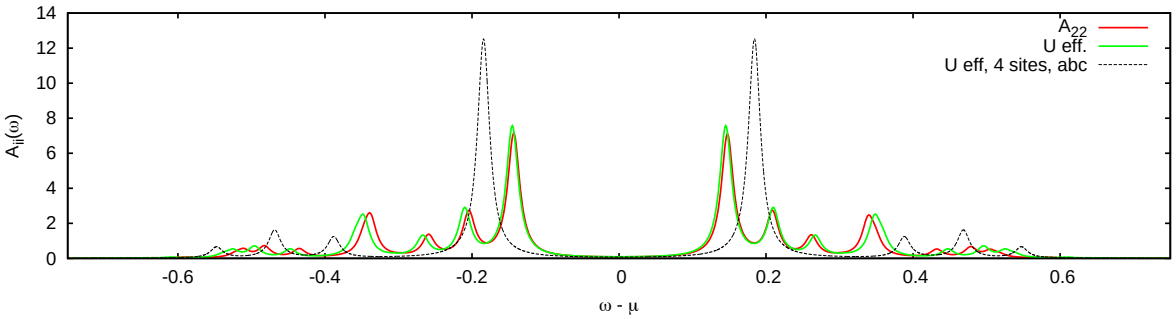


Figure 4.10.: Spectral function for parameter set (Sr), $\Delta = 1.4$ for a system of $L = 6$ sites and periodic boundary conditions. The single-band spectral function for the system with $L = 4$ sites is shown for comparison (dashed line).

calculations is indeed a finite size effect as it becomes significantly smaller when more sites and periodic boundary conditions are used. We also see that, fortunately, the static screening approach still accurately reproduces the Hubbard gap.

In conclusion, we have identified the parameter regime for which the single band description of the d-electrons is, at least partially, successful. The Hubbard U can be estimated to satisfactory accuracy using the static screening approach described in this section. However, the spectra are not reproduced perfectly, as there is some deviation among the minor peaks.

4.2. Instantaneous Screening Approximation

In this section, we formally introduce and describe an approximation that can be used to derive both the screening of the Hubbard U and the chemical potential. This formal approach will provide a solid foundation for the heuristic approach of the previous section and deepen our understanding of the physics involved.

Basically, we treat the renormalization in the full three-band Hilbert space by writing down a wavefunction in this space and then applying an approximation to it. An assumption already made in the preceding section was that the r-electrons are in their ground state with respect to the d-electrons. From the resulting ground state energies for $N_d = 0, 1, 2$, we could then obtain the effective on-site energy ε and the effective Hubbard U. We call this assumption the *instantaneous screening approximation* (ISA). It states that the r-electrons immediately adjust their state to that of the d-electrons. This approximation is justified when the r-electrons are much faster than the d-electrons, which is the case for $t_{13} \gg t_{22}$. This is similar to the Born-Oppenheimer approximation for electrons and nuclei. There, the atoms are assumed to be much slower than the electrons, which is motivated by their larger mass. Hence, the electrons instantaneously adjust to the given core coordinates. In Sec. 4.1, the ISA was already used in the static case to obtain an effective ε and U for the effective single-band system. Now, we will formalize this approximation by writing down the full wavefunction in the three-band Hilbert space.

4.2.1. Effective Hamiltonian in the ISA

Recall the decomposition of the full Hamiltonian H into Hamiltonians H_d, H_r and H_{rd} (Eqns. 3.2, 3.3 and 3.4). As they conserve the number of electrons in both the central and the outer bands, we can decompose the full Hilbert space $\mathcal{H}$ into a space for the d-electrons, $\mathcal{H}_d$, and a space for the r-electrons, $\mathcal{H}_r$, with $\mathcal{H} = \mathcal{H}_d \otimes \mathcal{H}_r$. Let $|\psi_n^d\rangle$ and $|\psi_m^r\rangle$ be complete sets of basis states for $\mathcal{H}_d$ and $\mathcal{H}_r$, respectively. Then we can write any state $|\Psi\rangle \in \mathcal{H}$ as

$$|\Psi\rangle = \sum_n \sum_m a_{nm} \, |\psi_n^d\rangle \otimes |\psi_m^r\rangle \, . \tag{4.10}$$

with coefficients a_{nm}. For the basis of $\mathcal{H}_d$, we choose the real space occupation number basis $|\mathbf{n_d}\rangle$. Then, $|\Psi\rangle$ is written as

$$|\Psi\rangle = \sum_{\mathbf{n_d}} \left(|\mathbf{n_d}\rangle \otimes \sum_m a_m(\mathbf{n_d}) \, |\psi_m^r\rangle \right) \tag{4.11}$$

with a new set of coefficients $a_m(\mathbf{n_d})$ where we replace the index n by $\mathbf{n_d}$. Let us consider at this moment the issue of antisymmetrization. In (4.10) and (4.11), the d- and r-electron states are antisymmetrized, but the product states are not. The question arises whether we have to antisymmetrize them as well. The general answer is yes, but because the Hamiltonian preserves N_d and N_r we are able to distinguish between the d- and r-electrons. This means that they can be treated as different particles. In fact, it does not

matter whether or not we antisymmetrize when dealing with operators that respect this separation. A formal proof is given in Appendix C.

We now subtly modify the wavefunction (4.11). Note that the basis set for $\mathcal{H}_r$ used in (4.11) to denote the r-electron wavefunction $|\psi^r\rangle$ need not be the same for each index n, or $\mathbf{n_d}$. This means that for each basis state of $\mathcal{H}_d$, we can use a different set of basis states for $\mathcal{H}_r$. This results in

$$|\Psi\rangle = \sum_{\mathbf{n_d}} |\mathbf{n_d}\rangle \otimes \sum_m a_m(\mathbf{n_d}) \, |\psi^r_m(\mathbf{n_d})\rangle \tag{4.12}$$

where now the r-electron basis states are parametrized with the d-electron configuration. The basis states of $\mathcal{H}_r$ are now parametrized with the configuration of the d-electrons. This is one step towards our goal of deriving an effective d-Hamiltonian in the occupation number basis of the d-electrons. For simplicity, we will omit the superscript r for the r-electron basis states from now on.

The full Schrödinger equation in this representation is

$$\begin{aligned} H\,|\Psi\rangle &= E\,|\Psi\rangle \\ (H_d + H_r + H_{rd}) \sum_{\mathbf{n_d}} \sum_m a_m(\mathbf{n_d}) \, |\mathbf{n_d}\rangle \otimes |\psi_m(\mathbf{n_d})\rangle &= E \sum_{\mathbf{n_d}} \sum_m a_m(\mathbf{n_d}) \, |\mathbf{n_d}\rangle \otimes |\psi_m(\mathbf{n_d})\rangle \,. \end{aligned} \tag{4.13}$$

We will now make a choice for the parametrized basis states of $\mathcal{H}_r$ that is well suited to the physical situation. The first step is to realize that the coupling Hamiltonian H_{rd} is diagonal in the configuration basis. Therefore, we can write

$$H_{rd} \, |\mathbf{n_d}\rangle \otimes |\psi_m(\mathbf{n_d})\rangle = |\mathbf{n_d}\rangle \otimes H_{rd}(\mathbf{n_d}) \, |\psi_m(\mathbf{n_d})\rangle \tag{4.14}$$

where $H_{rd}(\mathbf{n_d})$ is H_{rd} with the operators $n_{i,2\sigma}$ replaced by the corresponding number from $\mathbf{n_d}$. We can think of $H_{rd}(\mathbf{n_d})$ as an operator acting on $\mathcal{H}_r$ that is parametrized with a d-electron configuration. This is analogous to the treatment of atomic coordinates in the Born-Oppenheimer approximation, where the operators $\hat{\mathbf{R}}_\alpha$ are replaced by parameters $\mathbf{R}_\alpha$ (cf. Eq. 1.3 in Sec. 1.2.1). The next step is to choose as the parametrized basis states $|\psi_m(\mathbf{n_d})\rangle$ the eigenstates of $H_r + H_{rd}(\mathbf{n_d})$, that is,

$$(H_r + H_{rd}(\mathbf{n_d})) \, |\psi_m(\mathbf{n_d})\rangle = \mathcal{E}_m(\mathbf{n_d}) \, |\psi_m(\mathbf{n_d})\rangle \,. \tag{4.15}$$

Because we are interested in a Schrödinger equation for the d-Electrons, we multiply (4.13) from the left with $\langle\psi_n(\mathbf{n'_d})| \otimes \langle\mathbf{n'_d}|$. Using (4.15) we arrive at

$$a_n(\mathbf{n'_d})\mathcal{E}_n(\mathbf{n'_d}) + \sum_{\mathbf{n_d},m} a_m(\mathbf{n_d}) \, \langle\psi_n(\mathbf{n'_d})|\psi_m(\mathbf{n_d})\rangle \, \langle\mathbf{n'_d}|H_d|\mathbf{n_d}\rangle = a_n(\mathbf{n'_d})E. \tag{4.16}$$

This is a matrix equation for the coefficients $a_m(\mathbf{n_d})$. We have $\dim \mathcal{H}_{13}$ values for m and $\dim \mathcal{H}_2$ values for $\mathbf{n_d}$, so the total number of coefficients is $\dim \mathcal{H}_{13} \cdot \dim \mathcal{H}_2$. Although this is already significantly smaller than the full Hilbert space $\mathcal{H}$, we will now introduce the crucial approximation.

The simplification consists in introducing a cut-off for n. In the simplest case, only the ground-state of $H_r + H_{rd}$ is used. That is, we remove all r-electron degrees of freedom. For every configuration of the d-electrons, there is only one r-electron state retained. Eq. (4.16) then becomes

$$\boxed{a(\mathbf{n_d})\mathcal{E}_0(\mathbf{n_d}) + \sum_{\mathbf{n'_d}} a(\mathbf{n'_d}) \langle\psi_0(\mathbf{n_d})|\psi_0(\mathbf{n'_d})\rangle \langle\mathbf{n'_d}|H_d|\mathbf{n_d}\rangle = a(\mathbf{n_d})E.} \tag{4.17}$$

This is the matrix form of an effective Hamiltonian for the d-electrons. In this approximation, which we will call the *Instantaneous Screening Approximation* (ISA), the r-electrons instantaneously adjust to a change in the d-configuration. This has two effects, which are related to the diagonal and off-diagonal matrix elements of (4.17).

- The diagonal elements are readily obtained as

$$\mathcal{E}_0(\mathbf{n_d}) + \underbrace{\langle\psi_0(\mathbf{n_d})|\psi_0(\mathbf{n_d})\rangle}_{=1} \langle\mathbf{n_d}|H|\mathbf{n_d}\rangle = \mathcal{E}_0(\mathbf{n_d}) + U_{22} \cdot D(\mathbf{n_d}), \tag{4.18}$$

  where $D(\mathbf{n_d})$ is the number of doubly occupied sites in the configuration $\mathbf{n_d}$. This is precisely the static screening as discussed in Sec. 4.1. It is due to the energy term $\mathcal{E}_0(\mathbf{n_d})$, which introduces for each lattice site i an energy term that depends on the number of d-electrons on that site. Note that the screening is additive rather than multiplicative, so the r-electrons determine the absolute rather than the relative change of U_{22}. In our model, where the r-electron problem is independent for each lattice site, we can write the energy as a sum over the separate lattice sites, $\mathcal{E}_0(\mathbf{n_d}) = \sum_i \mathcal{E}_0^i(n_{d,i})$ with

$$\mathcal{E}_0^i(n_{d,i}) = \varepsilon_0 + n_{d,i}\varepsilon_i + (U - U_{22})n_{d,i,\uparrow}n_{d,i,\downarrow}. \tag{4.19}$$

  This corresponds to the terms in Fig. 4.2, only that here the energy due to the bare interaction of the d-electrons, U_{22}, was excluded from $\mathcal{E}_0^i(2)$ and included individually in (4.18).

- The off-diagonal elements are

$$\langle\psi_0(\mathbf{n_d})|\psi_0(\mathbf{n'_d})\rangle \langle\mathbf{n'_d}|H|\mathbf{n_d}\rangle \tag{4.20}$$

  The last factor of (4.20) is the normal off-diagonal matrix element of H_d. In our case it is, up to the appropriate Fermi sign, the hopping matrix element t_{22} when $\mathbf{n_d}$ and $\mathbf{n'_d}$ are connected by a single next-neighbor hopping process. We see that within the ISA, the effective Hamiltonian contains hopping factors that are reduced by the overlap matrix element

$$\langle\psi_0(\mathbf{n_d})|\psi_0(\mathbf{n'_d})\rangle\,. \tag{4.21}$$

  This hopping reduction was not accounted for in the previous, heuristic approach to screening. It explains why the minor peaks in the spectral functions of the effective single-band system were slightly too far out from the center.

In our model, where the r-electrons on each site are independent of the electrons at other sites, we can provide a particularly simple description of the hopping reduction. Let $|0\rangle$ denote $|\psi_0(n_d = 0)\rangle$ for a one-site system and define $|1\rangle$ and $|2\rangle$ analogously. These are the states depicted in Fig. 4.2. Then there are four distinct possible hopping processes, which are shown in Fig. 4.11. Thus, given the overlaps $\langle 0|1\rangle$ and $\langle 1|2\rangle$, we know how to modify the hopping matrix element t_{22} for each hopping process. Note, however, that

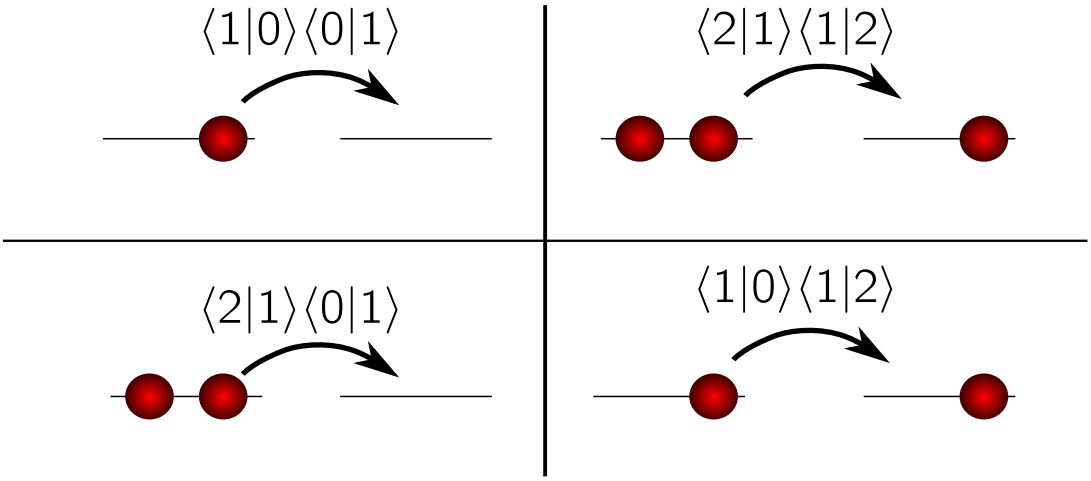


Figure 4.11.: Hopping processes in the instantaneous screening approximation and the associated reduction factor.

hopping now becomes a configuration dependent process. It cannot be written as a simple term of the form $-tc_i^\dagger c_j$.

The ISA is a variational approach, because the product states $|\mathbf{n_d}\rangle\,|\psi_0(\mathbf{n_d})\rangle$ form the basis of a subspace $\mathcal{H}_{\mathrm{ISA}}$ of the full Hilbert space $\mathcal{H}$. As the states of $\mathcal{H}_{\mathrm{ISA}}$ are completely determined by the configuration of the central band, $|\mathbf{n_d}\rangle$, there exists a one-to-one mapping g from $\mathcal{H}_{\mathrm{ISA}}$ to $\mathcal{H}_d$ which is readily obtained as

$$g : \mathcal{H}_{\mathrm{ISA}} \longrightarrow \mathcal{H}_d : |\mathbf{n_d}\rangle\,|\psi_0(\mathbf{n_d})\rangle \longrightarrow |\mathbf{n_d}\rangle_{\mathrm{eff}}\,. \tag{4.22}$$

From this, it immediately follows that $\dim(\mathcal{H}_{\mathrm{ISA}}) = \dim(\mathcal{H}_d)$. Thus, the ISA significantly reduces the size of the relevant Hilbert space (cf. Tab. 2.1). This is equivalent to integrating out all r-electron degrees of freedom.

Implementation

In principle, the ISA retains the full Hamiltonian yet restricts the full Hilbert space $\mathcal{H}$ to the subspace $\mathcal{H}_{\mathrm{ISA}}$ defined by the $|\psi_0(\mathbf{n_d})\rangle$. While this is conceptually correct, our aim was to derive an effective single-band Hamiltonian. To this end, we use the linear one-to-one mapping g (4.22). This means that we have to find the effective Hamiltonian H_{eff} that satisfies the following commutative diagram.

$$\begin{array}{ccc} |\Psi\rangle \in \mathcal{H}_{\text{ISA}} & \xrightarrow{\quad g \quad} & |\Psi_{\text{eff}}\rangle \in \mathcal{H}_d \\ \Big\downarrow{\scriptstyle PHP} & & \Big\downarrow{\scriptstyle H_{\text{eff}}} \\ |\tilde{\Psi}\rangle \in \mathcal{H}_{\text{ISA}} & \xrightarrow{\quad g \quad} & |\tilde{\Psi}_{\text{eff}}\rangle \in \mathcal{H}_d \end{array}$$

In the end, both approaches yield the same effective Hamiltonian, because from the commutative diagram we expect that

$$\langle \mathbf{n}'_\mathbf{d}| \langle \psi_0(\mathbf{n}'_\mathbf{d})|H|\psi_0(\mathbf{n}_\mathbf{d})\rangle |\mathbf{n}_\mathbf{d}\rangle = \langle \mathbf{n}'_\mathbf{d}|H_{\text{eff}}|\mathbf{n}_\mathbf{d}\rangle \,, \tag{4.23}$$

that is, the matrix elements of the Hamiltonian in $\mathcal{H}_{\text{ISA}}$ and those of the effective single-band Hamiltonian H_{eff} are the same. The important difference is that when using the full Hamiltonian in $\mathcal{H}_{\text{ISA}}$, we still have to deal with the large many-body vectors $|\mathbf{n}_\mathbf{d}\rangle |\psi_0(\mathbf{n}_\mathbf{d})\rangle$ of $\mathcal{H}$, while using the effective single-band Hamiltonian H_{eff} allows us to perform all calculations with the much smaller vectors $|\mathbf{n}_\mathbf{d}\rangle$ of $\mathcal{H}_d$. We have implemented both approaches for testing purposes but practice only the latter.

Effective Hamiltonian in Second Quantization

We can formally derive the connection between the effective Hamiltonian H_{eff} and the full Hamiltonian H by introducing a set of transformed particle operators, $\tilde{c}_{i\sigma}$ and $\tilde{c}^\dagger_{i\sigma}$. Apart from removing or inserting a d-electron, these operators also adjust the state of the r-electrons. A formal way to describe this is via their action on the basis states of $\mathcal{H}_{\text{ISA}}$, which is

$$\begin{aligned} \tilde{c}_{i\sigma} |\mathbf{n}_\mathbf{d}\rangle |\psi_0(\mathbf{n}_\mathbf{d})\rangle &= n_{d,i,\sigma}(-1)^{\pi(i,\sigma;\mathbf{n}_\mathbf{d})} |\mathbf{n}'_\mathbf{d}\rangle |\psi_0(\mathbf{n}'_\mathbf{d})\rangle \\ \tilde{c}^\dagger_{i\sigma} |\mathbf{n}_\mathbf{d}\rangle |\psi_0(\mathbf{n}_\mathbf{d})\rangle &= (1-n_{d,i,\sigma})(-1)^{\pi(i,\sigma;\mathbf{n}_\mathbf{d})} |\mathbf{n}'_\mathbf{d}\rangle |\psi_0(\mathbf{n}'_\mathbf{d})\rangle \end{aligned} \tag{4.24}$$

Here, $n_{d,i,\sigma} = \hat{n}_{d,i,\sigma} |\mathbf{n}_\mathbf{d}\rangle$. This prefactor ensures that we cannot remove an electron where there is none, while the prefactor $(1 - n_{d,i,\sigma})$ enforces the Pauli exclusion principle. We define $\pi(i, \sigma, \mathbf{n}_\mathbf{d})$ as the permutation that moves the particle operator $\tilde{c}^{(\dagger)}$ to its canonical position within the product of creation operators for $|\mathbf{n}_\mathbf{d}\rangle$ (cf. Sec. 2.3.1). Then $(-1)^\pi$ is defined as the sign of the permutation, that is, 1 for an even permutation and -1 for an odd permutation.

Let us now express these modified particle operators in terms of the original particle operators. Because how the modified operators act on the r-electrons depends on the configuration of the d-electrons, we cannot expect them to be a simple linear combination of the original operators. We start by writing down an expression that achieves the same as (4.24) and then simplify it. First, let

$$\hat{\psi}^\dagger_d(\mathbf{n}_\mathbf{d}) = \prod_\sigma \prod_i \left(c^\dagger_{d,i,\sigma}\right)^{n_{d,i,\sigma}} \tag{4.25}$$

be the operator that creates $|\mathbf{n}_\mathbf{d}\rangle$ when applied to the vacuum and let $\hat{\psi}^\dagger_0(\mathbf{n}_\mathbf{d})$ be the operator that creates $|\psi_0(\mathbf{n}_\mathbf{d})\rangle$. A formal definition is given in Appendix D.1. Second, let

$$\hat{P}_d(\mathbf{n_d}) = \prod_{i:\text{occ}} \hat{n}_i \prod_{i:\text{unocc}} (1 - \hat{n}_i) \tag{4.26}$$

be the operator that projects a state onto $|\mathbf{n_d}\rangle$. We call these operators *generalized particle operators*. With them we can write

$$\tilde{c}_{i\sigma} = \sum_{\mathbf{n_d}, n_{i\sigma}=1} (-1)^{\pi(i,\sigma;\mathbf{n_d})} \hat{\psi}_d^\dagger(\mathbf{n'_d}) \hat{\psi}_0^\dagger(\mathbf{n'_d}) \hat{\psi}_0(\mathbf{n_d}) \hat{\psi}_d(\mathbf{n_d}) \hat{P}_d(\mathbf{n_d}). \tag{4.27}$$

Let us examine the different parts of (4.27). To make the discussion easier to follow, we show how the different parts act on the ISA Wavefunction. The following equation is best left from right to left.

$$\underbrace{\hat{\psi}_d^\dagger(\mathbf{n'_d})\hat{\psi}_0^\dagger(\mathbf{n'_d})}_{\alpha_{\mathbf{n_d}}|\mathbf{n'_d}\rangle|\psi_0(\mathbf{n_d})\rangle} \underbrace{\hat{\psi}_0(\mathbf{n_d})\hat{\psi}_d(\mathbf{n_d})}_{\alpha_{\mathbf{n_d}}|0\rangle\langle 0|} \underbrace{\hat{P}_d(\mathbf{n_d})}_{\alpha_{\mathbf{n_d}}|\mathbf{n_d}\rangle|\psi_0(\mathbf{n_d})\rangle} \sum_{\alpha_{\mathbf{n_d}}} |\mathbf{n_d}\rangle\,|\psi_0(\mathbf{n_d})\rangle \tag{4.28}$$

First, we single out a particular configuration of the d-electrons using $\hat{P}_d$. We then remove all electrons, both d and r. Next, we reinsert the electrons in the new configuration $\mathbf{n'_d}$ that results from $\mathbf{n_d}$ by removing the electron with spin σ from site i. Note that in (4.28), no Fermi sign occurs. However, when an ordinary particle operator $c_{i,\sigma}^{(\dagger)}$ is applied to a state $|\mathbf{n_d}\rangle$, a Fermi sign occurs. Therefore, we explicitly include in our definition of $\tilde{c}_{i,\sigma}^{(\dagger)}$.

It is easy to see that the operator $\tilde{c}_{i\sigma}$, when defined as in (4.27) satisfies the definition (4.24). We now turn to the simplification of (4.27). The four generalized particle operators can be rearranged to

$$\hat{\psi}_d^\dagger(\mathbf{n'_d})\hat{\psi}_d(\mathbf{n_d})\hat{\psi}_0^\dagger(\mathbf{n'_d})\hat{\psi}_0(\mathbf{n_d})$$

with no Fermi sign occurring as the number of r-electrons remains unchanged. We can simplify the product of the d-particle operators by observing that the configurations $\mathbf{n_d}$ and $\mathbf{n'_d}$ differ only in that $\mathbf{n_d}$ contains an electron with spin σ at site i while $\mathbf{n'_d}$ does not. From this, it follows that

$$c_{d,i,\sigma}\hat{\psi}_d^\dagger(\mathbf{n_d}) = (-1)^{\pi(i,\sigma;\mathbf{n_d})}\hat{\psi}_d^\dagger(\mathbf{n'_d})(1 - \hat{n}_{d,i,\sigma}) \tag{4.29}$$

where we use the fermionic anti-commutation rules to first move $c_{d,i,\sigma}$ up to its counterpart and then to eliminate it. By removing $c_{d,i,\sigma}^\dagger$ from $\hat{\psi}_d^\dagger(\mathbf{n_d})$, we arrive at $\hat{\psi}_d^\dagger(\mathbf{n'_d})$. With

$$\hat{n}_{d,i,\sigma}\hat{\psi}_d(\mathbf{n_d}) \ldots \hat{P}_d(\mathbf{n_d}) = 0,$$

we can use (4.29) to rewrite (4.27):

$$\begin{aligned}
\tilde{c}_{i\sigma} &= \sum_{\mathbf{n_d}, n_{i\sigma}=1} c_{d,i,\sigma}\hat{\psi}_d^\dagger(\mathbf{n_d})\hat{\psi}_d(\mathbf{n_d})\hat{\psi}_0^\dagger(\mathbf{n'_d})\hat{\psi}_0(\mathbf{n_d})\hat{P}_d(\mathbf{n_d}) \\
&= c_{d,i,\sigma} \sum_{\mathbf{n_d}, n_{i\sigma}=1} \hat{\psi}_0^\dagger(\mathbf{n'_d})\hat{\psi}_0(\mathbf{n_d})\hat{P}_d(\mathbf{n_d}) \\
&= c_{d,i,\sigma}\hat{\Psi}_-(i,\sigma)
\end{aligned} \tag{4.30}$$

Let us again see how the modified operator in this form acts on the ISA basis states.

$$\begin{aligned}\tilde{c}_{i,\sigma}\left|\mathbf{n_d}\right\rangle\left|\psi_0(\mathbf{n_d})\right\rangle &= c_{d,i,\sigma}\hat{\psi}_0^\dagger(\mathbf{n_d'})\hat{\psi}_0(\mathbf{n_d})\hat{\psi}_d^\dagger(\mathbf{n_d})\hat{\psi}_0^\dagger(\mathbf{n_d})\left|0\right\rangle \\ &= c_{d,i,\sigma}\hat{\psi}_d^\dagger(\mathbf{n_d})\hat{\psi}_0^\dagger(\mathbf{n_d'})\left|0\right\rangle \\ &= (-1)^{\pi(i,\sigma;\mathbf{n_d})}\left|\mathbf{n_d'}\right\rangle\left|\psi_0(\mathbf{n_d'})\right\rangle .\end{aligned} \tag{4.31}$$

This was for one ISA basis state; the projector $\hat{P}_d(\mathbf{n_d})$ ensures that this result also holds for linear combinations of the basis states. This projection is necessary, because the change in the r-electron wavefunction due to the creation or annihilation of a d-electron is, in general, configuration dependent and therefore has to be done separately for each d-electron configuration $\mathbf{n_d}$. In the special case where the r-electron state factors into on-site states that are independent from each other and where the change in the r-electron state is linear in the number of d-electrons per site, the projector can be omitted and we just have

$$\tilde{c}_{i,\sigma}^{(\dagger)} = c_{i,\sigma}^{(dg)}\psi_{0,i}^\dagger(n_{d,i}\pm 1)\psi_{0,i}(n_{d,i}). \tag{4.32}$$

Putting everything together, we see that $\Psi_-(i,\sigma)$ is the operator that realizes the change in the r-electron state due to the removal of a d-electron in state $i\sigma$. Analogously, we write $\Psi_+(i,\sigma)$ when a d-electron is added. One expects that $\Psi_+(i,\sigma)$ is the adjoint of $\Psi_-(i,\sigma)$. This is formally shown in Appendix D.2. In Appendix D.2.1, we also show that the anti-commutation relations of $\tilde{c}_{i\sigma}$ and $\tilde{c}_{i\sigma}^\dagger$ are just the same as for normal fermionic particle operators.

Now that we have properly defined the modified particle operators $\tilde{c}$ and $\tilde{c}^\dagger$, we can turn our attention to the effective Hamiltonian. Let H_{ISA} be H_d with the particle operators $c_i^{(\dagger)}$ replaced by the modified particle operators $\tilde{c}_i^{(\dagger)}$, i.e.,

$$H_{\mathrm{ISA}} = -t\sum_{\substack{\sigma\\ i,j:\mathrm{nn}}}\tilde{c}_{i\sigma}^\dagger\tilde{c}_{j\sigma} + U\sum_i \tilde{n}_{i\uparrow}\tilde{n}_{i\downarrow}, \tag{4.33}$$

which is just H in $\mathcal{H}_{\mathrm{ISA}}$ (cf. Eq.4.17) with the diagonal energy term $\mathcal{E}_0(\mathbf{n_d})$ absorbed into U_{22} to yield the effective Hubbard U. We attempt to rewrite this in terms of the original particle operators. For this, we look at the operator product $\tilde{c}_i^\dagger\tilde{c}_j$. For $i=j$, we have

$$\begin{aligned}\tilde{c}_{i\sigma}^\dagger\tilde{c}_{i\sigma} &= c_{i\sigma}^\dagger c_{i\sigma}\sum_{\mathbf{n_d}}\hat{\psi}_0^\dagger(\mathbf{n_d})\hat{\psi}_0(\mathbf{n_d})\hat{P}_d(\mathbf{n_d}) \\ &= c_{i\sigma}^\dagger c_{i\sigma}\sum_{\mathbf{n_d}}\hat{P}_d(\mathbf{n_d}) \\ &= c_{i\sigma}^\dagger c_{i\sigma}\end{aligned} \tag{4.34}$$

so $\tilde{n}_{i\sigma} = n_{i\sigma}$. This is the second-quantization equivalent to the observation in (4.18) that the r-electron wavefunctions for the diagonal elements of the effective Hamiltonian have overlap 1. Similar manipulations as used in the derivation of (4.30) lead to

$$\tilde{c}_{i\sigma}^\dagger\tilde{c}_{j\sigma} = c_{i\sigma}^\dagger c_{j\sigma}\sum_{\substack{\mathbf{n_d}\\ n_{i\sigma}=0\\ n_{j\sigma}=1}}\hat{\psi}_0^\dagger(\mathbf{n_d}^{j\sigma\to i\sigma})\hat{\psi}_0(\mathbf{n_d})\hat{P}_d(\mathbf{n_d}). \tag{4.35}$$

Here, $\mathbf{n_d}^{j\sigma\to i\sigma}$ is the configuration that arises from $\mathbf{n_d}$ when an electron with spin σ moves from site j to site i. This is just the second-quantization version of the hopping reduction arising from the overlap matrix element $\langle\psi(\mathbf{n_d})|\psi(\mathbf{n_d})\rangle$ in (4.20). If we now integrate out the r-electrons, we arrive at H_{eff}, a Hubbard Hamiltonian for the d-electrons alone, of the form

$$H_{\text{eff}} = -t_{\text{eff}} \sum_{\substack{\sigma \\ i,j:\text{nn}}} c^\dagger_{i\sigma} c_{j\sigma} + U_{\text{eff}} \sum_i n_{i\uparrow} n_{i\downarrow} \tag{4.36}$$

with renormalized Hubbard parameters.

4.2.2. Green's Functions in the Instantaneous Screening Approximation

The physical idea behind the ISA is that the r-electrons instantaneously adjust to the electronic configuration of the d-electrons. In other words, the r-electrons are much faster than the d-electrons. While this holds for the dynamics described by our Hamiltonian H, it is not clear whether this still holds for the dynamics of electron emission or inverse electron emission. Because the Green's Functions are somewhat related to these two processes, we must spend some time deliberating how these processes are treated in the instantaneous screening approximation. In the next paragraphs we will only talk about electron insertion, because the arguments are the same for electron emission.

There are two limiting cases for the process of electron emission. When the process occurs on a time scale much faster than all other dynamics of the system, an electron is removed from the central band without the r-electrons adjusting to this change. This limiting case is called the *sudden approximation*. When the process occurs on a time scale much slower than the dynamics of the r-electrons, removing an electron from the central band results in the r-electrons immediately adjusting to the new electron configuration in the d-band. This is called the *adiabatic approximation*. Let us go into the details of these two cases.

In frequency space, the diagonal Green's Function for an electron on site i and spin σ in the central band is given by

$$\begin{aligned} G_{d,i,\sigma}(\omega) = \quad & \langle\Psi_0|c^\dagger_{d,i,\sigma}\left(\omega - H + E_0 + i\eta\right)^{-1} c_{d,i,\sigma}|\Psi_0\rangle \\ & + \langle\Psi_0|c_{d,i,\sigma}\left(\omega + H - E_0 + i\eta\right)^{-1} c^\dagger_{d,i,\sigma}|\Psi_0\rangle\,. \end{aligned} \tag{4.37}$$

Because the resolvent can be expanded in a power series of H, we study the repeated application of H to $c^\dagger_{d,i,\sigma}|\Psi\rangle$. In the instantaneous screening approximation, we use PHP instead of H, where P is the projector onto $\mathcal{H}_{\text{ISA}}$. Depending on how we implement the insertion (or emission) of the d-electrons, using the project affects the sum-rule for the spectral weights (Eq. 4.8). An arbitrary state of $\mathcal{H}_{\text{ISA}}$ can be written as a linear combination of the basis states

$$|\Psi\rangle = \sum_{\mathbf{n_d}} \alpha(\mathbf{n_d})\,|\mathbf{n_d}\rangle\,|\psi_0(\mathbf{n_d})\rangle\,. \tag{4.38}$$

The adiabatic approximation is realized when, in the definition of the Green's Function, (4.37) we replace the normal particle operators with the modified particle operators (4.24).

Note that the operators are *not* replaced within the Hamiltonian H. From (4.24) it follows that $\tilde{c}^{(\dagger)} |\Psi\rangle \in \mathcal{H}_{\mathrm{ISA}}^{N\pm1}$ for each $|\Psi\rangle \in \mathcal{H}_{\mathrm{ISA}}$. For that reason, using these modified operators for calculating the Green's Function leads to

$$\tilde{c}^{\dagger}_{d,i,\sigma} |\Psi\rangle = \sum_{\mathbf{n_d}} \alpha(\mathbf{n_d})(1 - n_{d,i,\sigma})(-1)^{\pi(i,\sigma;\mathbf{n_d})} |\mathbf{n'_d}\rangle |\psi_0(\mathbf{n'_d})\rangle . \tag{4.39}$$

The resulting state is part of the reduced Hilbert space and consequently, applying the projector P does not change it. As a consequence, the sum rule for the spectral weights is not affected.

For the sudden approximation, we use the unmodified particle operators, as then only the d-electrons are affected while the r-electrons do not relax. Let us see how applying the creation operator for a d-electron acts on an ISA-state.

$$\begin{aligned} c^{\dagger}_{d,i,\sigma} |\Psi\rangle &= \sum_{\mathbf{n_d}} \alpha(\mathbf{n_d}) c^{\dagger}_{d,i,\sigma} |\mathbf{n_d}\rangle |\psi_0(\mathbf{n_d})\rangle \\ &= \sum_{\mathbf{n_d}} \alpha(\mathbf{n_d})(1 - n_{d,i,\sigma})(-1)^{\pi(i,\sigma;\mathbf{n_d})} |\mathbf{n'_d}\rangle |\psi_0(\mathbf{n_d})\rangle . \end{aligned} \tag{4.40}$$

Again, $\mathbf{n'_d}$ is the state that results from $\mathbf{n_d}$ when a central band electron with spin σ is added to site i. This state is undefined should there already be an electron of that spin at that site, but this is taken care of by the factor $(1 - n_{d,i,\sigma})$. In general, the state described by (4.40) does not lie in the reduced Hilbert space $\mathcal{H}_{\mathrm{ISA}}$, since the state of the r-electrons is still the one corresponding to $\mathbf{n_d}$ while the state of the d-electrons changed to $\mathbf{n'_d}$. We now apply the projection operator P to that state. To this end, we expand the projection operator in our basis of $\mathcal{H}_{\mathrm{ISA}}$.

$$P = \sum_{\mathbf{n_d}} |\mathbf{n_d}\rangle |\psi_0(\mathbf{n_d})\rangle \langle\psi_0(\mathbf{n_d})| \langle\mathbf{n_d}| . \tag{4.41}$$

The result of applying P to (4.40) is then readily obtained as

$$\begin{aligned} P c^{\dagger}_{d,i,\sigma} |\Psi\rangle &= \sum_{\mathbf{n_d}} \alpha(\mathbf{n_d})(1 - n_{d,i,\sigma})(-1)^{\pi(i,\sigma;\mathbf{n_d})} P |\mathbf{n'_d}\rangle |\psi_0(\mathbf{n_d})\rangle \\ &= \sum_{\mathbf{n_d}} \alpha(\mathbf{n_d})(1 - n_{d,i,\sigma})(-1)^{\pi(i,\sigma;\mathbf{n_d})} \sum_{\tilde{\mathbf{n}}_d} |\tilde{\mathbf{n}}_d\rangle |\psi_0(\tilde{\mathbf{n}}_d)\rangle \langle\psi_0(\tilde{\mathbf{n}}_d)| \langle\tilde{\mathbf{n}}_d|\mathbf{n'_d}\rangle |\psi_0(\mathbf{n_d})\rangle \\ &= \sum_{\mathbf{n_d}} \alpha(\mathbf{n_d})(1 - n_{d,i,\sigma})(-1)^{\pi(i,\sigma;\mathbf{n_d})} |\mathbf{n'_d}\rangle |\psi_0(\mathbf{n'_d})\rangle \langle\psi_0(\mathbf{n'_d})|\psi_0(\mathbf{n_d})\rangle . \end{aligned} \tag{4.42}$$

We see that the sudden approximation involves a matrix element $\langle\psi_0(\mathbf{n'_d})|\psi_0(\mathbf{n_d})\rangle$. Because of this additional matrix element, the sum rule for the spectral function does no longer hold, i.e., $||A(\omega)||_1 < 1$.

Eqns. (4.39) and (4.42) describe how the adiabatic approximation works for vectors that are part of the full Hilbert space We now turn our attention to the effective single-band Hilbert space $\mathcal{H}_d$ and investigate how the Green's Function calculated with H_{eff} in $\mathcal{H}_d$

relates to the Green's Function calculated with PHP in $\mathcal{H}_{\mathrm{ISA}}$. It suffices to look at the basis states $|\mathbf{n_d}\rangle$ of $\mathcal{H}_d$. We have

$$c^\dagger_{d,i,\sigma}\,|\mathbf{n_d}\rangle = (1-n_{d,i,\sigma})(-1)^{\pi(i,\sigma;\mathbf{n_d})}\,|\mathbf{n'_d}\rangle \xrightarrow{g^{-1}} (1-n_{d,i,\sigma})(-1)^{\pi(i,\sigma;\mathbf{n_d})}\,|\mathbf{n'_d}\rangle\,|\psi_0(\mathbf{n'_d})\rangle\,. \quad (4.43)$$

Hence, we see that using the normal particle operators in $\mathcal{H}_d$ corresponds to using the adiabatic approximation in $\mathcal{H}_{\mathrm{ISA}}$. To implement the sudden approximation, we have to modify these operators by explicitly including the overlap matrix elements between different r-electron states.

Results

Before we continue with our theoretical considerations, let us briefly pause and look at an example of the hopping reduction in action. Fig. 4.12 shows the spectral function obtained with the ISA when the full many-body nature of the hopping reduction, i.e., the dependence of the hopping amplitude on the number of d-electrons on that site, is taken into account. For comparison, we also show the spectral function for static screening of

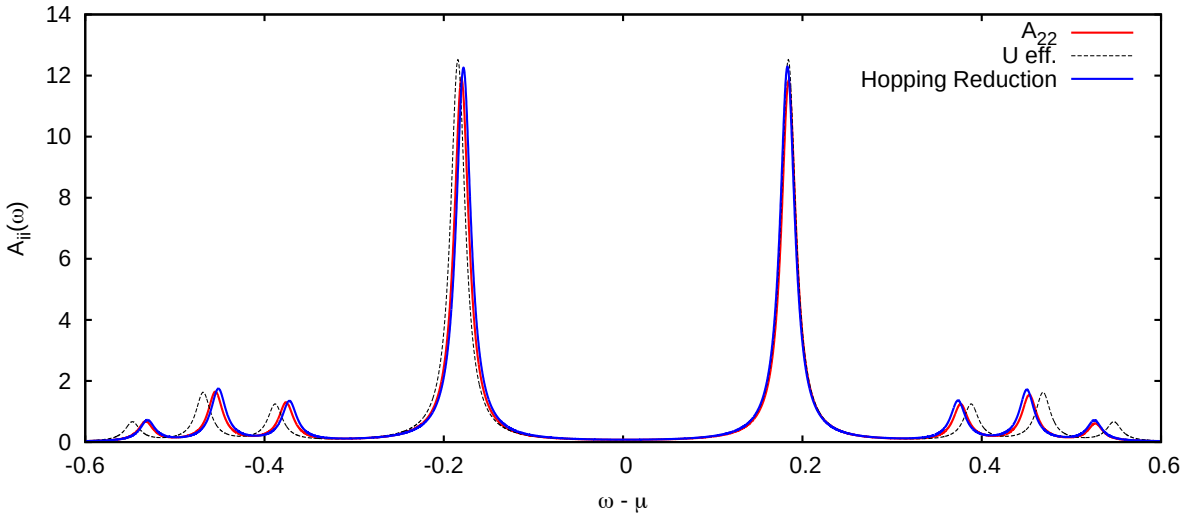


Figure 4.12.: Spectral function obtained within the ISA for the full many-body hopping reduction (blue). The spectral function of the d-electrons is shown in red. For comparison, the spectral function obtained in the static screening approach is shown as the dotted line. Parameters: Sr from Tab. 4.2 at $\Delta = 1.4$.

U only. We clearly see that the ISA gives significantly better results. In addition to the Mott gap, the minor peaks agree much better with those of the full system. Instead of overestimating t_{22}, we now slightly underestimate it. This is a result of our calculation of the ψ_0 in the atomic limit. In reality, the central band electrons are spread out among multiple sites due to the hopping. Informally, we can say that the r-electrons on one site anticipate the hopping of d-electrons from adjacent sites. Formally, we can argue with the variational principle. Currently, we assume that $|i\rangle$ is the correct state of the

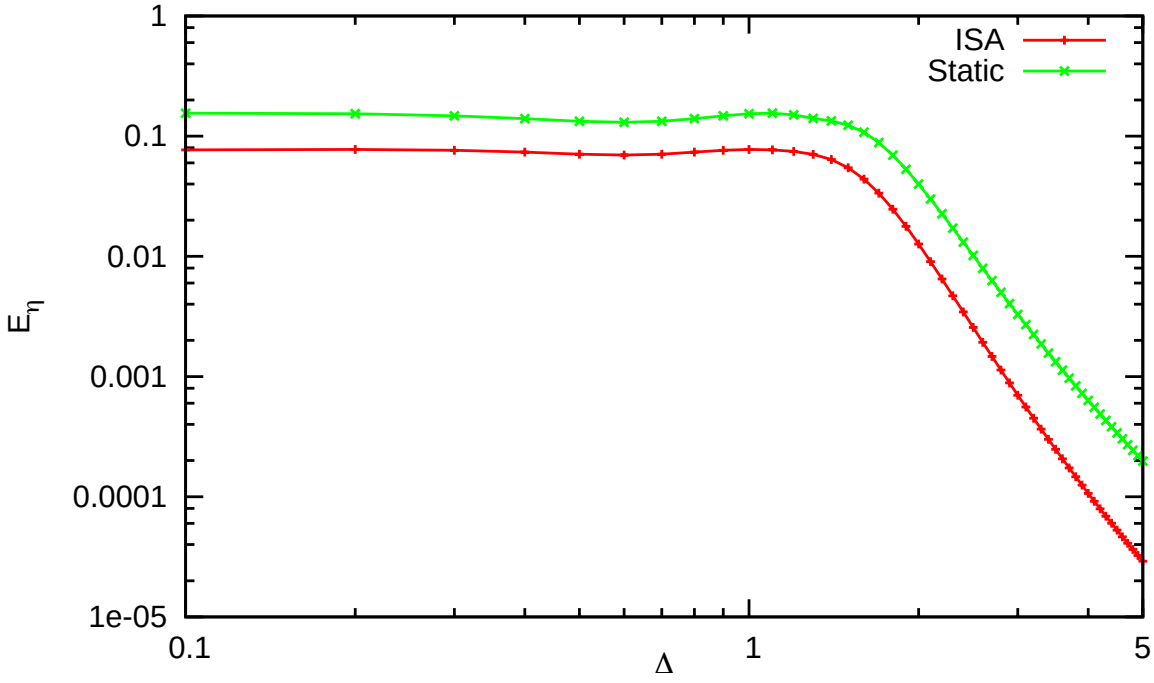


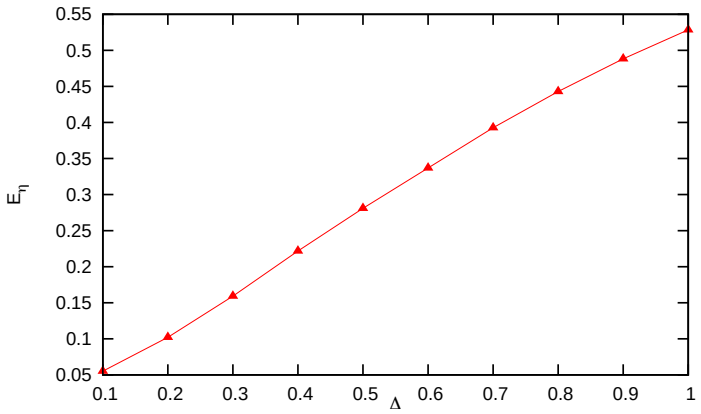
Figure 4.13.: Error measure E_η with $\eta = 0.01$ for A_{22} of the full system with A_{ISA} (red curve) and A_{static} (green curve). We see that the ISA yields clearly superior results when compared to the static screening without hopping reduction. Parameters Sr from Tab. 4.2.

r-electrons of a site with i electrons in the ground state. We replace $|0\rangle$ with $|0\rangle + \varepsilon\,|1\rangle$ and $|2\rangle$ with $|2\rangle + \varepsilon\,|1\rangle$. Then, we can find an ε that minimizes the resulting ground state energy, which will improve upon our approximation. These states then have more overlap with $|1\rangle$, so that the hopping matrix elements become slightly larger again. This comes at the expense of the screening of U, as can easily be seen in the limit of unscreened electrons, where all $|i\rangle$ are the same. Here, the hopping is unreduced, but the Coulomb interaction is unscreened as well. For an alternative explanation for the overestimation of the hopping reduction, we note that the ISA is a special case of our decoupling of the r- and d-electrons where we made a cut-off of $n = 0$ in Eq. 4.16. By including higher lying excitations of the r-electrons, we will get a higher effective hopping matrix element as well.

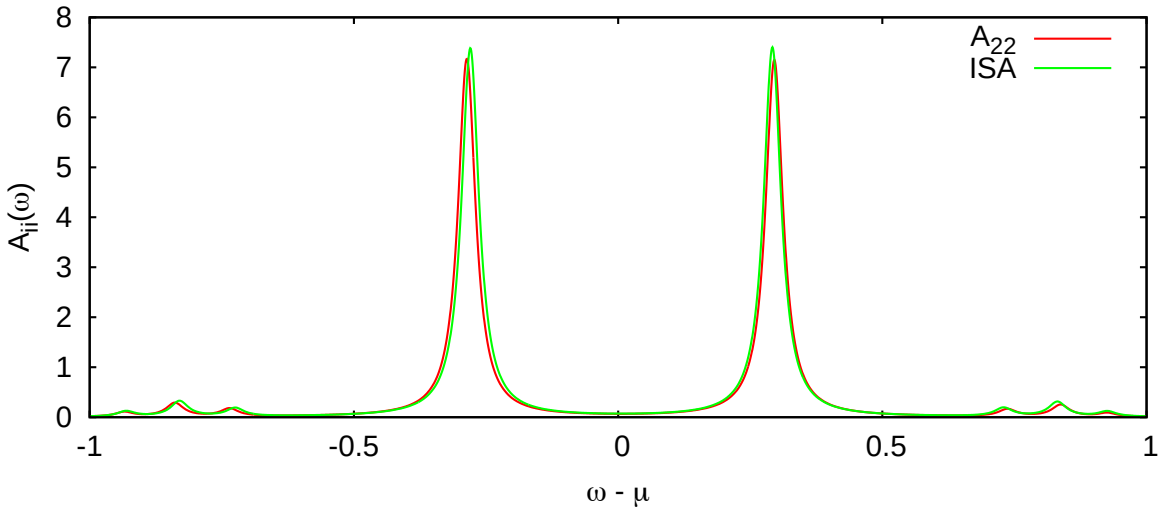
Now that we have seen the ISA work for one specific value of Δ, we use our error measure to evaluate the quality of the ISA for the entire range of band separations. For parameter set (Sr) of Tab. 4.2, we compute the spectral function and compare it to that of the full system. The results are shown in Fig. 4.13. We clearly see that the error measure obtained within the ISA is at least a factor of 2 lower than that obtained from using an effective U only, without taking the hopping reduction into account.

We expect that for faster d-electrons, the ISA becomes worse, although it will still perform better than using only an effective U and keeping t at its bare value. In conclusion, the overestimation of the hopping reduction should increase with increasing t. We show

this in Fig. 4.14.



(a) Error measure, $\eta = 0.02$



(b) $t_{22} = -0.2$.

Figure 4.14.: Error measure and example of a spectral function for larger hopping t_{22} in parameter set (Sr). We observe that the error measure rises with rising t_{22}, as expected.

4.2.3. Effective Hubbard-Model

Although the ISA works remarkably well, it is not exactly what we are looking for. Our aim is to obtain an effective Hubbard model for the d-electrons. Currently, we have a Hubbard U and, if the total energy is of interest, an orbital energy ε. What we do not have, however, is a single value for the hopping amplitude t. In the ISA, the instantaneous reaction of the r-electrons to the d-electrons introduces matrix elements that reduce the hopping amplitude depending on how many d-electrons are on the sites. This makes

hopping a many-body effect, which is formalized in replacing the normal particle operators, $c_i^{(\dagger)}$, by the modified particle operators $\tilde{c}_i^{(\dagger)}$. In this section, we explore additional approximations on top of the ISA that give us a single hopping amplitude t.

Linear Response

When we assume the correlation between the r-electrons and the d-electrons to be small, the change in the r-wavefunction should be linear in the number of d-electrons on the site. We therefore write

$$|1\rangle = \frac{|0\rangle + |\varepsilon\rangle}{\sqrt{1+\varepsilon^2}} \tag{4.44}$$

$$|2\rangle = \frac{|0\rangle + 2\,|\varepsilon\rangle}{\sqrt{1+4\varepsilon^2}} \tag{4.45}$$

with $|\varepsilon\rangle$ being a small perturbation with norm ε and with $\langle 0|\varepsilon\rangle = 0$. It can be derived, for example, in second order perturbation theory, which yields a linear correction to wavefunctions. Let $|k^0\rangle$ be the eigenstates of the unperturbed system, with $|0^0\rangle$ being the ground state, and let E_k^0 be the corresponding eigenenergies.) Then, the correction to the ground state is

$$|\varepsilon\rangle = \sum_{k\neq 0} \frac{\langle k^0|V|0^0\rangle}{E_0^0 - E_k^0}\,|k^0\rangle\,. \tag{4.46}$$

Thus, it is *linear* in the perturbation potential V. Here, the perturbation are the terms $n_d U_{12}$ and $n_d U_{23}$ due charges on the central orbital.

Up to second order in ε, we have

$$\langle 0|1\rangle = \langle 1|2\rangle = 1 - \frac{\varepsilon^2}{2} + \mathcal{O}(\varepsilon^3). \tag{4.47}$$

We see that for small ε the overlap matrix elements are the same. Consequently, the hopping reduction is independent of the number of d-electrons on each site. We can just replace the bare value of t_{22} with $t_{22}\,\langle 0|1\rangle^2$ and obtain essentially the same results as in the ISA. The remaining question is under which conditions the response of the r-wavefunction is indeed linear. Therefore, we first look beyond the linear response approximation.

Linear Combination of Overlaps

If the overlap matrix elements $\langle 0|1\rangle$ and $\langle 1|2\rangle$ are different, we can still arrive at an effective single band Hubbard model with a single hopping parameter t by forming an appropriate combination of the matrix elements. The different matrix elements are associated with the four processes in Fig. 4.11. To combine them into a single hopping reduction factor, we use the probabilities E, P and D, which were introduced in Sec. 4.1.2, and weigh the hopping factor of every process with the probability of that process. The unnormalized

hopping probabilities are

$$\begin{aligned} P(10 \to 01) &= P \cdot E \\ P(11 \to 02) &= P^2 \\ P(20 \to 11) &= D \cdot E \\ P(21 \to 12) &= D \cdot P. \end{aligned} \tag{4.48}$$

The normalization factor is $(1-E)(1-D)$, as can be seen by

$$\frac{PE + P^2 + DE + DP}{(1-E)(1-D)} = \frac{P(E+P+D) + DE}{1-E-D+ED} = \frac{P+DE}{P+ED} = 1. \tag{4.49}$$

Note that in Eqs. 4.48 we take the probabilities of different sites as being uncorrelated. This approximation is also made in the GWA mentioned in Sec. 4.1.2. As we will see, the resulting spectral functions are still in very good agreement with the full spectrum.

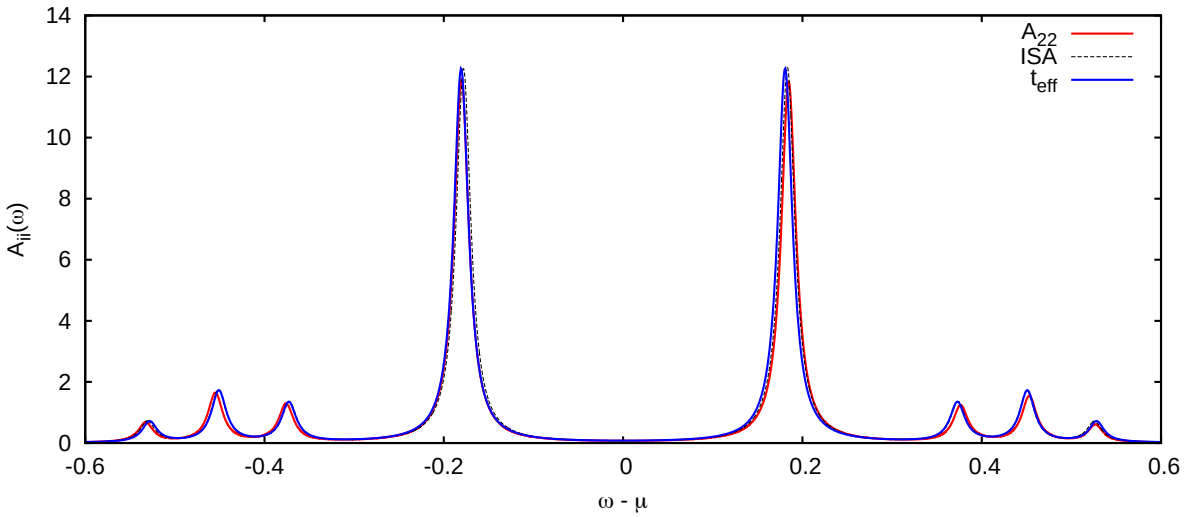


Figure 4.15.: Spectral function of the single band model with both an effective U and an effective t (blue curve) compared to the spectral function of the full system (red curve) and of the full ISA (dotted). Parameters from Tab.4.2 (Sr) with $\Delta = 1.4$.

In Fig. 4.15 we show the resulting spectral function for our usual example set of parameters. There is almost no visible deviation from the spectral function obtained with the full ISA, so we can assume that for the given parameters the linear response approximation holds. In this regime, we also find that there is no difference between the sudden approximation and the adiabatic approximation, because the additional matrix elements introduced by the sudden approximation are all the same and thus only give rise to a global factor that slightly reduces the spectral function.

Let us briefly talk about the method of averaging the different hopping factors. The disadvantage of using the probabilities E, P, and D to weigh the hopping factors is that, in order to calculate these probabilities, we already have to solve the full three-band system, because we do not know the appropriate parameters of the single-band system

yet. As the main purpose of the effective single-band Hamiltonian is to avoid these costly calculations, we seem to have struck a dead end. A way out is to use some other averaging procedure. For the special case of a system at half filling, we have $E = D$. Inserting this into (4.48), we can recast the weighed average as follows.

$$t_{\mathrm{avg}} = \frac{t_{\mathrm{bare}}}{(1-D)^2}\left[2PD\underbrace{\frac{\langle 0|1\rangle^2 + \langle 1|2\rangle^2}{2}}_{A} + (P^2+D^2)\underbrace{\langle 0|1\rangle\,\langle 1|2\rangle}_{B}\right], \tag{4.50}$$

where A is the arithmetic mean of $\langle 0|1\rangle^2$ and $\langle 1|2\rangle^2$ while B is the geometric mean of them. Regardless of the specific weights that are used, the averaged factor will lie in the interval spanned by A and B. This interval, however, is very small, as the geometric and the arithmetic mean are close to each other. Instead of using the weights $2PD$ and $P^2 + D^2$, we can therefore approximate t_{avg} by the average of the two means,

$$t_{\mathrm{avg}'} = \frac{t_{\mathrm{bare}}}{2}\frac{\langle 0|1\rangle^2 + \langle 1|2\rangle^2}{2} + \frac{t_{\mathrm{bare}}}{2}\langle 0|1\rangle\,\langle 1|2\rangle\,. \tag{4.51}$$

Since $t_{\mathrm{avg}'}$ also lies in the interval spanned by A and B, it will be close to t_{avg}. In this approximation, all four hopping processes are weighed equally. This situation would occur for a system with $P = D = E = 1/3$. However, D is limited to $1/4$ in the uncorrelated case and will drop with increasing Hubbard U, so the justification for using $t_{\mathrm{avg}'}$ is unphysical. A better, and equally simple, approximation is to observe that in strongly correlated systems, D will be small. Therefore, we are justified in dropping term A from (4.50) altogether and use

$$t_{0112} = t_{\mathrm{bare}}\,\langle 0|1\rangle\,\langle 1|2\rangle\,. \tag{4.52}$$

The physical picture behind this approximation is that for strong repulsive on-site interaction, doubly occupied sites are not found in the ground state. Therefore, all hopping will be of the type $11 \to 02$, i.e., an electron hops from a singly occupied site to another singly occupied site, leading to an empty and a doubly occupied site. We have performed calculations for the usual example parameters (Tab. 4.2 (Sr, $\Delta = 1.4$) and found almost no difference between the different approximations used for t_{avg}. We therefore do not show the corresponding spectral functions. In Fig. 4.16, we show the deviation of the resulting single-band spectral functions from the full spectral function. We observe that the choice of the averaging method is completely irrelevant. For the parameters used, the resulting hopping reduction matrix elements are shown in Fig. 4.17(a). For the value of $\Delta = 1.4$, we have $\langle 0|1\rangle = 0.98$ and $\langle 1|2\rangle = 0.96$. They are only 0.02 apart. Therefore, the resulting values of t_{avg} for different averaging procedures differ by even less, so that this difference is insignificant.

Let us now try to obtain a more interesting situation, where the choice of how to obtain a single effective t actually matters. We will try to make the overlap between the $|i\rangle$ smaller. This requires a detailed analysis of how these overlap matrix elements depend on the other parameters. We can make use of the exact diagonalization of the on-site Hamiltonian that

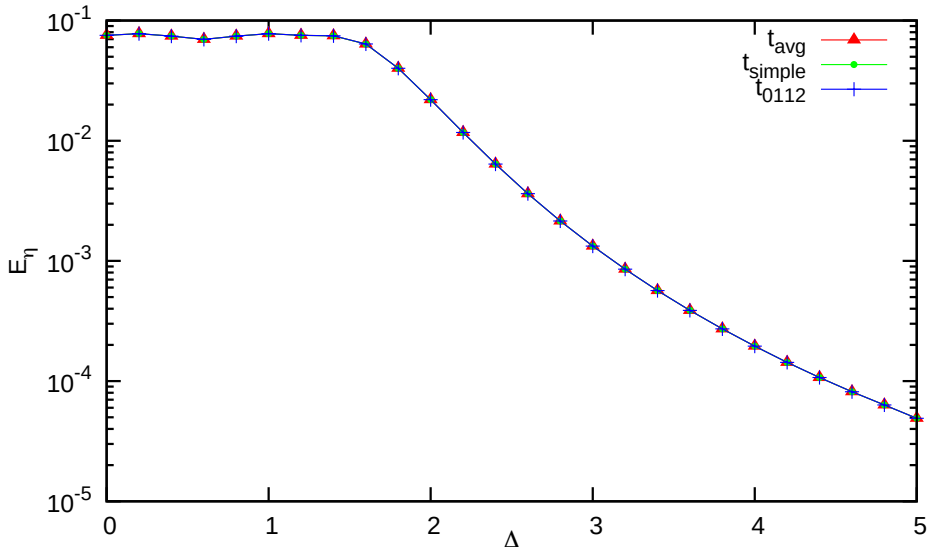


Figure 4.16.: Error measure for spectral functions with different averaging method for t. Red: Weighed average (Eq. 4.50). Green: Simple averaging (Eq. 4.51). Blue: Hopping process with single sites only (Eq. 4.52).

becomes possible when all but U_{12}, U_{22} and U_{33} are zero (cf. Appendix B). This gives us the ground state of the r-electrons for a given number of d-electrons as

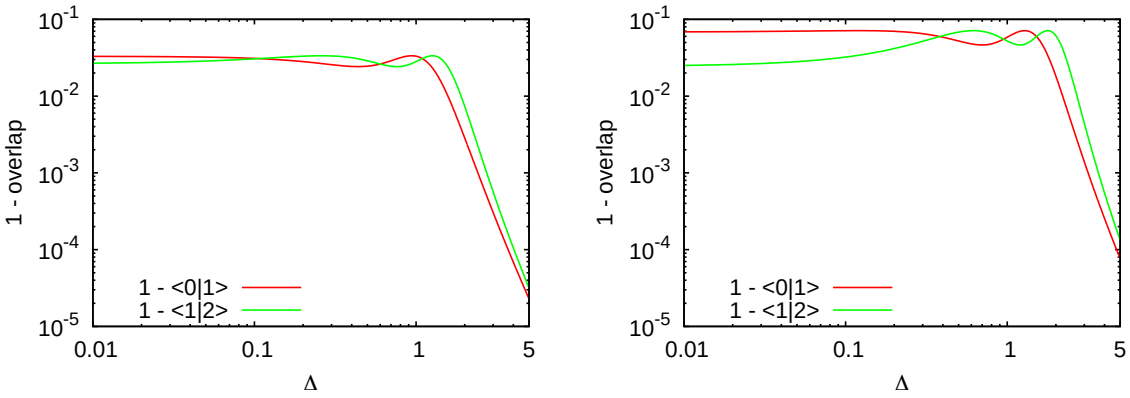
$$\frac{1}{\sqrt{A_i^2 + 4t_{13}^2}} \begin{pmatrix} -\frac{1}{2}\sqrt{A_i^2 + 4t_{13}^2} - \frac{1}{2}A_i \\ -\frac{1}{2}\sqrt{A_i^2 + 4t_{13}^2} + \frac{1}{2}A_i \\ t_{13} \\ t_{13} \end{pmatrix}, \tag{4.53}$$

cf. Eq. B.4. Here, A_i is a shorthand for $2\Delta - n_2(U_{12} - U_{23})$. We can use this ground state vector to calculate the overlap matrix elements $\langle 0|1\rangle$ and $\langle 1|2\rangle$, but, unfortunately, the resulting expressions become too complicated to evaluate. However, we can intuitively state that increasing $U_{12} - U_{23}$ and decreasing t_{13} might lead to a larger difference of the overlap matrix elements. Consulting the discussions about the model parameters in Sec. 3.3.1, we take a parameter region where the change δn_1 of the lower orbital's occupation is large. The parameters are shown in Tab. 4.3. We have verified that the orbital occupation behaves as intended for $n_d = 0, 1, 2$. The screening of t with these

Table 4.3.: Parameters for strong difference in the $|i\rangle$.

U_{11}	2.0	U_{12}	1.0	t_{22}	-0.1
U_{22}	0.6	U_{23}	0.0	t_{13}	-0.5
U_{33}	0.2				

parameters is shown in Fig. 4.17(b) For $\Delta = 1.8$, we have $|0|1\rangle = 0.967$ and $|1|2\rangle = 0.923$, which amounts to a difference of almost 5%. For this parameter set, we calculate the



(a) Params. (Sr) (b) Params. Tab. 4.3

Figure 4.17.: For the hopping reduction matrix elements $\langle 0|1\rangle$ and $\langle 1|2\rangle$ the difference to the unscreened limit, 1.

spectral functions with our various approaches for obtaining a screened U and t. The result is shown in Fig. 4.18. Although the difference between the hopping reduction matrix elements is quite large, a difference among the spectral functions with an effective t is barely visible. We conclude that an effective t averaged over the different hopping reduction factors suffices in all the situations we have considered.

4.2.4. Doped Systems

So far, only systems at half filling were studied. It is, however, interesting to also look at doped systems with additional electrons or holes. First, a remark about the computational problems arising for doped system. We consider here a system with an additional electron. As in the half-filled case we always had an even number of electrons because we restricted ourselves to an even number of sites, we now have an odd number of electrons. We have discussed in Sec. 3.3.2 that in this case, the ground state is always degenerate. To lift the degeneracy, we shift the orbital energy of one of the d-orbitals by a small amount $\varepsilon = 10^{-4}$. This may appear quite large, actually, but degeneracy was still present for smaller values. In our calculations, we also found that convergence to the true ground state was much more problematic. The cut-off precision for the ground state energy had to be raised from 10^{-12}, which yields reasonable results for systems at half filling, to 10^{-16}.

When the system is doped with one additional electron of spin σ, there are two different spectra we can obtain, $A_{ii,\sigma}$ and $A_{ii,\bar{\sigma}}$. It is therefore common to calculate both and take their average, $A_{ii} = 1/2(A_{ii,\sigma} + A_{ii,\bar{\sigma}}$. For the parameter set Sr, we show the resulting deviation of our various single-band approaches from the full system's spectrum in Fig. 4.19. The first thing to note is that the screening of t is very important, as the error for the ISA with unscreened U is almost an order of magnitude higher than the errors for the various hopping reduction approaches. Again, we see that using an effective t is not problematic and that the fully configuration dependent hopping is not needed to

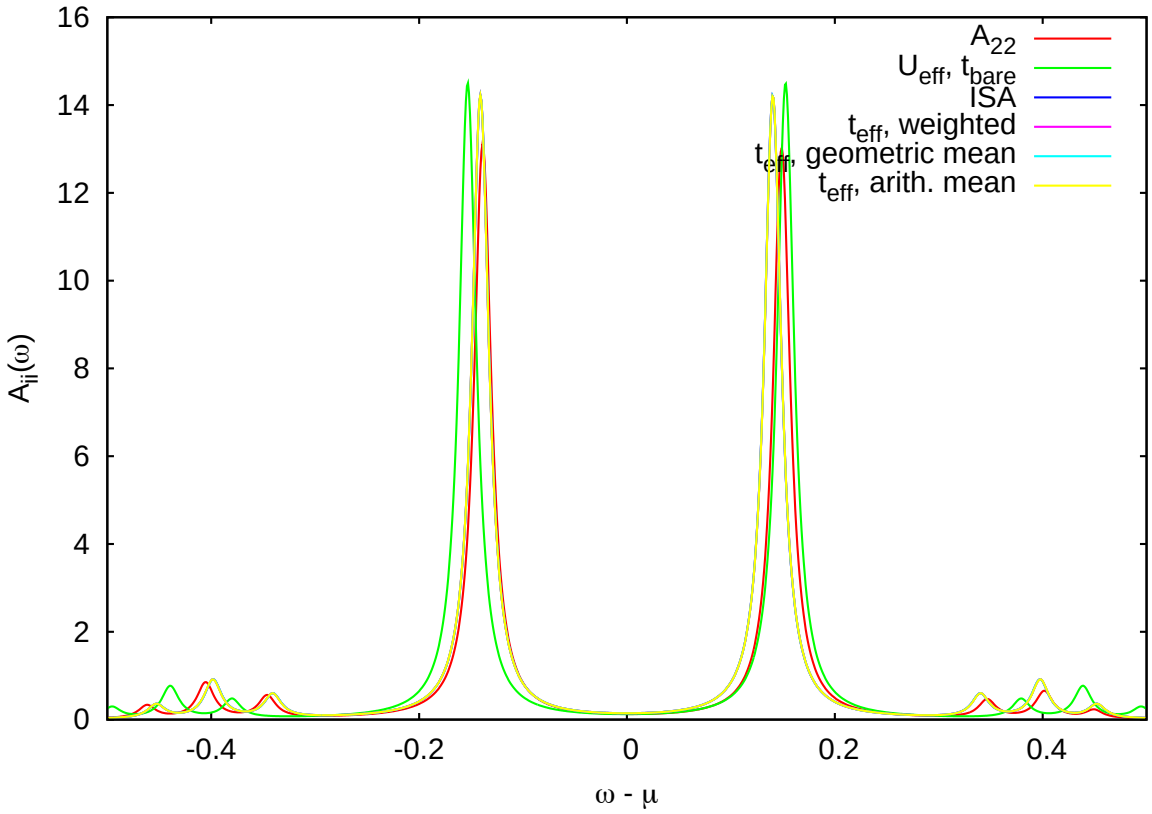


Figure 4.18.: Spectral functions for the parameter set Tab. 4.3

yield accurate results.

For $\Delta = 1.4$, we show the resulting spectral functions in Fig. 4.20(a). The chemical potential can be obtained graphically by plotting the PES and IPES part of the spectrum separately. Then, μ lies above the last peak of the photoemission spectrum and below the first peak of the inverse photoemission spectrum (cf. Fig. 4.20(b)). In agreement with the low error measure observed in Fig. 4.19, the effective single band model provides an accurate description of the d-band.

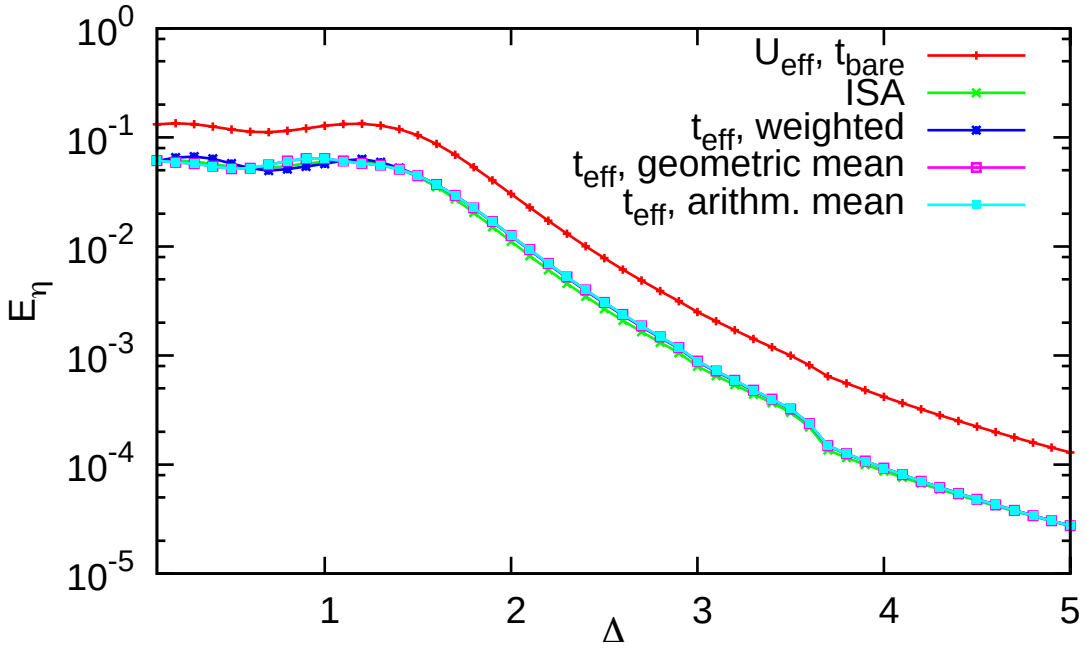
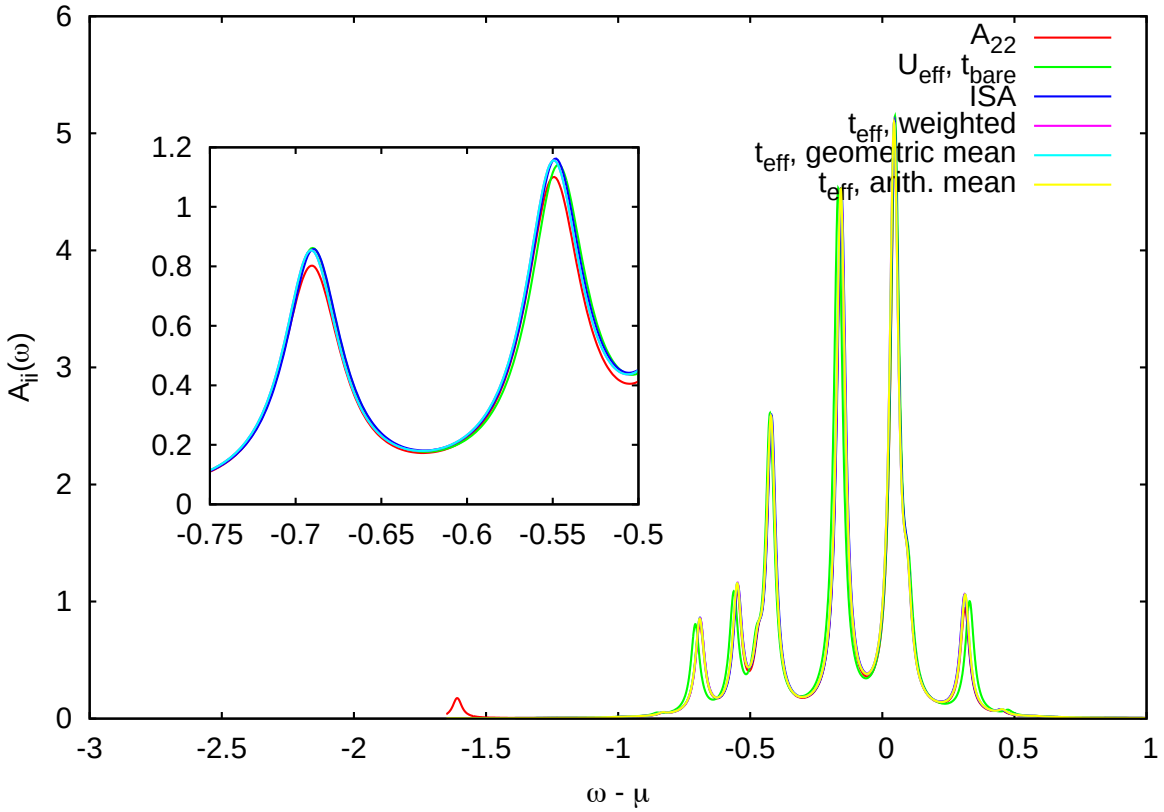
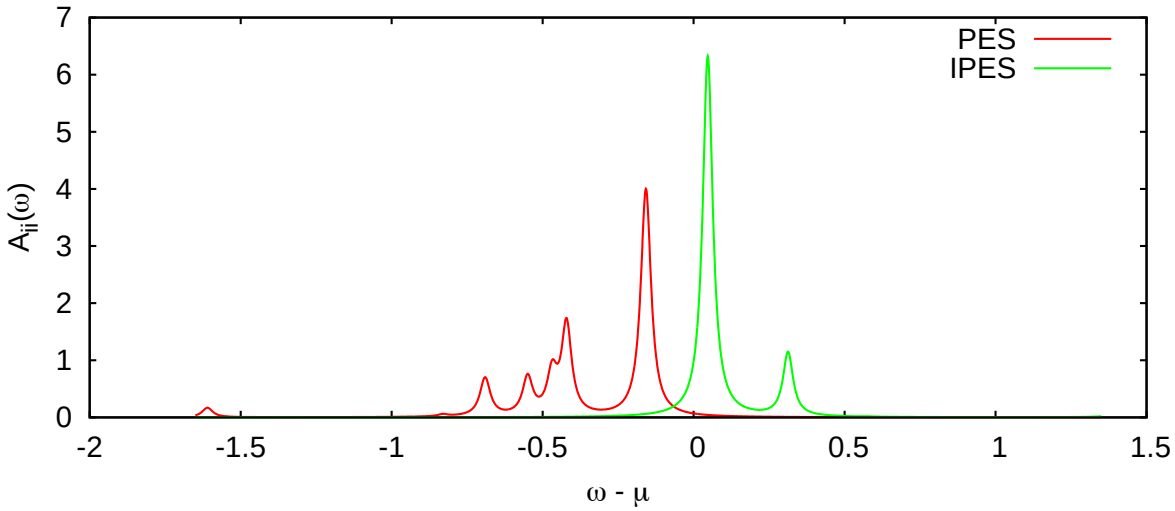


Figure 4.19.: Error measure ($\eta = 0.02$) for ISA without hopping reduction (red), full configuration dependent ISA (green), effective t with weighted mean (blue), effective t with $t = t_{\text{bare}} \langle 0|1\rangle \langle 1|2\rangle$ (magenta) and effective t with $t = t_{\text{bare}}/2(\langle 0|1\rangle^2 + \langle 1|2\rangle^2)$ (cyan).



(a) Spectral Functions



(b) Full system, PES and IPES separately

Figure 4.20.: Spectral functions for a system with an additional electron, averaged over both spin types. Parameter set (Sr) with $\Delta = 1.4$. η is 0.02.

5. Random Phase Approximation

So far, our method of reducing the full system to an effective one includes an exact diagonalization for the r-electron subspace. In realistic situations, however, the r-electrons themselves will be a complex many-body system that calls for an approximate treatment. One such approximation is the Random Phase Approximation (RPA). Constrained RPA, or c-RPA, has been introduced by F. Aryasetiawan to obtain a frequency dependent, screened Hubbard U [23, 50, 51]. We will now apply the c-RPA to our model and study how accurate its results are. We begin with an introduction to the formalism and then compare it to the ISA.

5.1. Screening of U in c-RPA

We will briefly restate the main idea of the constrained RPA as introduced by F. Aryasetiawan. The fully screened Coulomb interaction is given as

$$W = [1 - vP]^{-1} v \tag{5.1}$$

where v is the unscreened Coulomb interaction and P is the polarization. In RPA, the Polarization is given by

$$P(\mathbf{r}, \mathbf{r}'; \omega) = \sum_{i}^{\text{occ}} \sum_{j}^{\text{unocc}} \psi_i(\mathbf{r}) \psi_i^*(\mathbf{r}') \psi_j^*(\mathbf{r}) \psi_i^*(\mathbf{r}') \left(\frac{1}{\omega - \varepsilon_j + \varepsilon_i + i\eta} - \frac{1}{\omega + \varepsilon_j - \varepsilon_i - i\eta} \right), \tag{5.2}$$

where the ψ_i are single-particle eigenfunctions corresponding to the system's band structure with the ε_i as the corresponding eigenenergies. The approach of Aryasetiawan is to split the full polarization P into a part for the strongly correlated d-electrons and a part for the rest,

$$P = P_d + P_r, \tag{5.3}$$

where P_d includes only transitions between the d-electrons, while P_r contains all the rest. Note that this rest is *not* restricted to transitions between r-electrons, but also includes those between r- and d-electrons. In [50], Aryasetiawan and his collaborators showed that this splitting allows us to rewrite (5.1) into

$$\begin{aligned} W &= [1 - vP_r - vP_d]^{-1} v \\ &= \left[(1 - vP_r) \left\{ 1 - (1 - vP_r)^{-1} vP_d \right\} \right]^{-1} v \\ &= \left\{ 1 - (1 - vP_r)^{-1} vP_d \right\}^{-1} (1 - vP_r)^{-1} v \\ &= [1 - UP_d]^{-1} U. \end{aligned} \tag{5.4}$$

We can interpret

$$U(\omega) = [1 - vP_r(\omega)]^{-1} v. \tag{5.5}$$

as an effective interaction among the d-electrons. $U(\omega)$ is the *partially* screened Coulomb interaction, where screening is only due to processes involving r-electrons as well. To obtain the *fully* screened Coulomb interaction, processes in the d-band have to be taken into account as well. From (5.5), we see that we retain the fully screened Coulomb interaction, within the RPA, from

$$W = (1 - U(\omega)P_d(\omega))^{-1} U(\omega). \tag{5.6}$$

By integrating out the r-electron degrees of freedom, the interaction of the d-electrons thus becomes frequency-dependent. While this can be taken into account by using an effective action instead of a Hamiltonian, we are interested in obtaining a single, static Hubbard U. In our case, where we assume the dynamics of the r-electrons to be much faster than the dynamics of the d-electrons, we just take the static limit, $U = U(\omega = 0)$. This means that we neglect any retardation effects of the r-electrons, which is in the same spirit of the ISA.

5.2. Renormalization of the Hubbard U

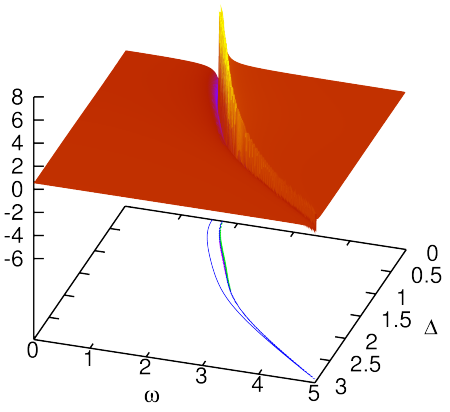
For our system, the quantities v, P, P_r and $U(\omega)$ are matrices of size $N \times N$ where N is the number of orbitals. Lacking any long-range interaction, however, these matrices are block-diagonal with local blocks of size 3×3. One block of the unscreened Coulomb interaction v has the form

$$\begin{pmatrix} U_{11} & U_{12} & 0 \\ U_{12} & U_{22} & U_{23} \\ 0 & U_{23} & U_{33} \end{pmatrix} \tag{5.7}$$

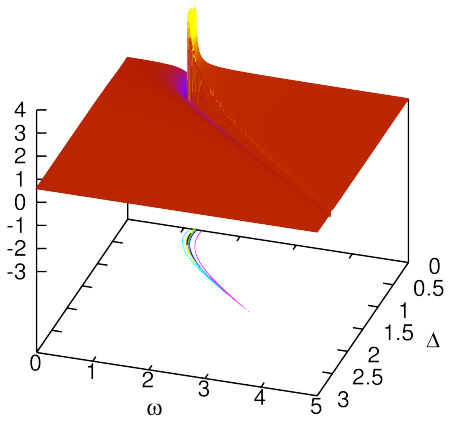
To compute P_r, we need the Bloch wavefunctions of the r-electrons. The d-electron states will not contribute to P_r, as in our given model they have no overlap with the r-electron states. Thus, we solve the on-site Hamiltonian for the r-electrons using the Hartree approximation. The resulting matrix is then plugged into (5.5) to yield the screened matrix $\tilde{U}$.

The resulting screened $U(\omega)$ is shown for parameter set (Sr) of Tab. 4.2 in Fig. 5.1(a). We observe a singularity that is due to the denominators in Eq. 5.2. For the non-interacting system, we would expect these singularities at $\omega = 2\sqrt{\Delta^2 + t_{13}^2}$. In the interacting case, the position will become lower in ω, as U_{11} will raise the energy of the lower orbital. Still, we see that the singularity occurs at larger values of ω when Δ is increased and that the weight of the poles decreases with Δ. This is because with increasing Δ, the overlap between the states of the lower and the upper r-band becomes smaller. We pick out the curve for $\Delta = 1.4$ in Fig. 5.1(c). The singularity lies at approximately $\omega = 2.6$.

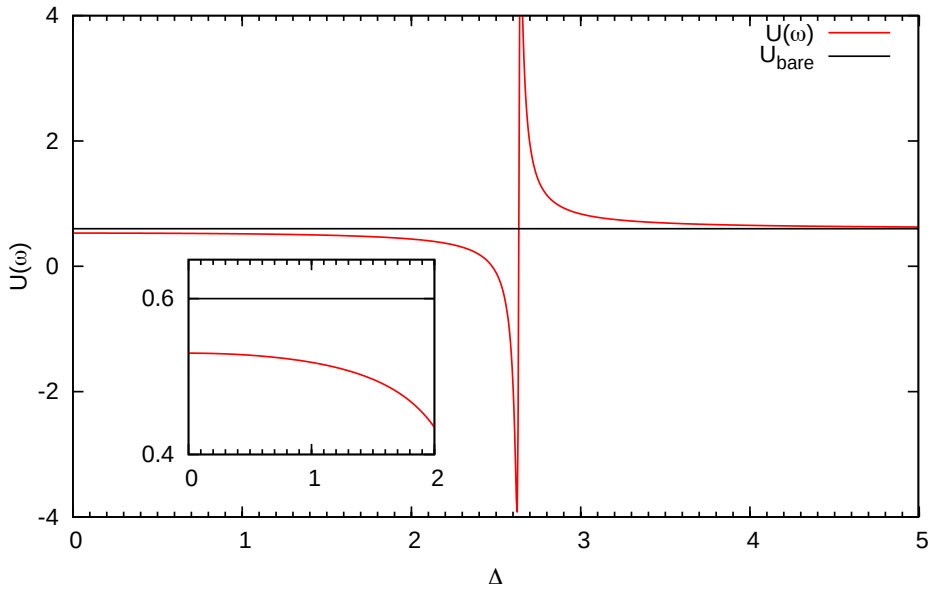
Let us now compare the single effective Hubbard U obtained for $U = U(\omega = 0)$ with that obtained by the ISA (Sec. 4.1). We do this for all four parameter sets of Tab. 4.2 and plot the relative strength of the screening in Fig. 5.2. For the parameter sets WrMcSd



(a) $U(\omega)$. Params: Sr



(b) $U(\omega)$. Parameters: WrMcSd



(c) Sr, $\Delta = 1.4$

Figure 5.1.: Frequency dependent effective U obtained from c-RPA. Parameters (Sr) from Tab. 4.2.

and WrWcSd, which show a discontinuity in the renormalization due to electrons leaving the d-band (cf. Sec.4.1), we also plot the value of U that is obtained when n_d is held fixed at the values 0,1, and 2. We call the U that is obtained this way U_{eff} and the U obtained from the static approach of Sec.4.1 U'_{eff}. In the case of weakly interacting r-electrons, the screening obtained by ISA and c-RPA is practically the same, except, of course, for the

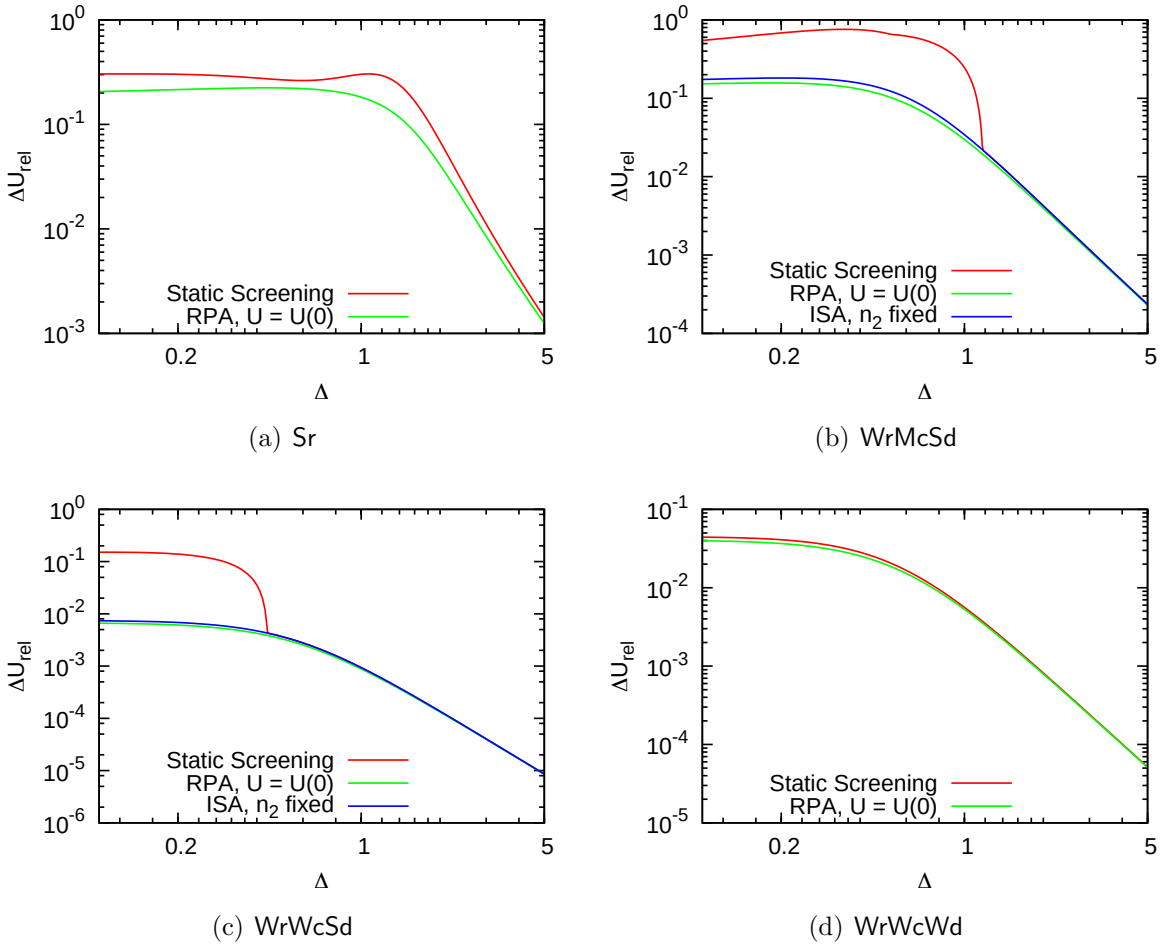


Figure 5.2.: Relative screening strength for ISA (red) and the static limit ($\omega = 0$) of c-RPA (green). For the two parameter sets WrMcSd and WrWcSd, where electrons leave the d-band, we also show U_{eff}.

case where electrons leave the d-band. For strongly interacting r-electrons (Fig. 5.2(a)), however, screening is overestimated by c-RPA. Using the error measure, we could confirm that the values of U obtained by the ISA yield the correct Mott gap (data not shown). This is because the eigenstates of the r-electrons are solved exactly in our approach, while a Hartree approximation is used in c-RPA. Therefore, we can expect the energies obtained within the ISA to be more accurate. This also explains why the difference between ISA and c-RPA vanishes for smaller interactions U_{11} and U_{33}.

Influence of the Resonance

It has been argued in [50] that the resonance found in the frequency dependent effective Coulomb interaction would have a large impact on the static U. This conflicts with our claim made in the introduction that higher lying excitations play only a minor role in

describing the low-energy physics of a system. In the previous calculations, we have seen that using the value of $U = U(\omega = 0)$ yields satisfactory results. We also see that only in the very vicinity of the resonance does the value of $U(\omega)$ deviate significantly from its static limit. An averaging process over a low frequency range will therefore yield values very close to $U(0)$ unless the resonance itself comes close to 0. From Fig. 5.1(b), however, we see that not even for parameters where the system becomes instable ($\Delta = 1$, for example) does the peak approach 0 but rather stays finite. In other words, the resonance never comes close enough to $\omega = 0$ to have significant influence on the low-frequency characteristics of the system. Thus, using the static limit of $U(\omega)$ always suffices in the cases we have considered.

Screening in the Unstable Regime

In Sec. 4.1, we showed as an example of unphysical behavior the spectral function for parameter set (WrMcSd) of Tab. 4.2 with $\Delta = 1.0$. There, we experienced poor agreement between the full spectral function and that of the effective single-band system (Fig. 4.8(d)).

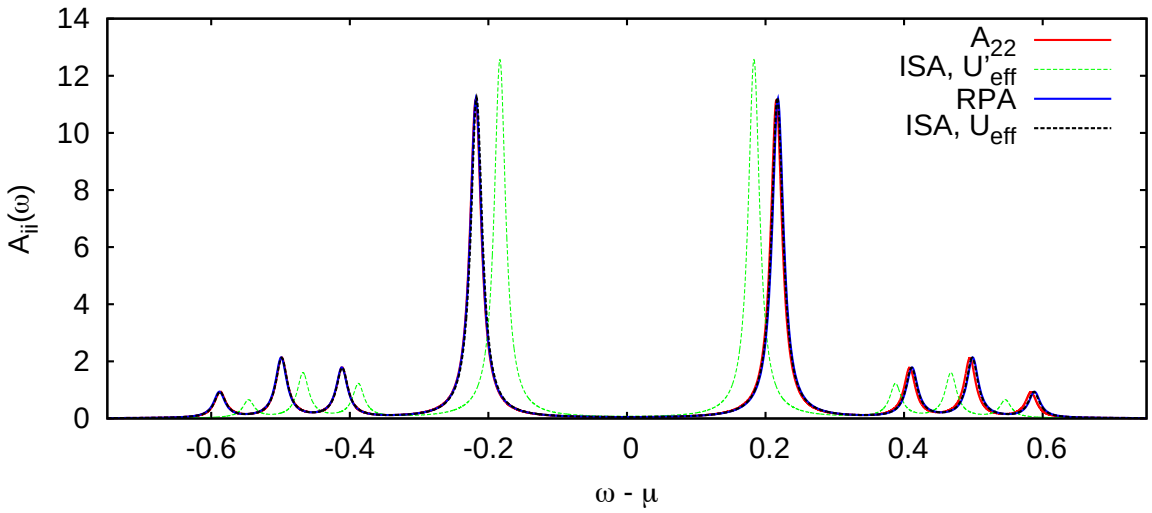


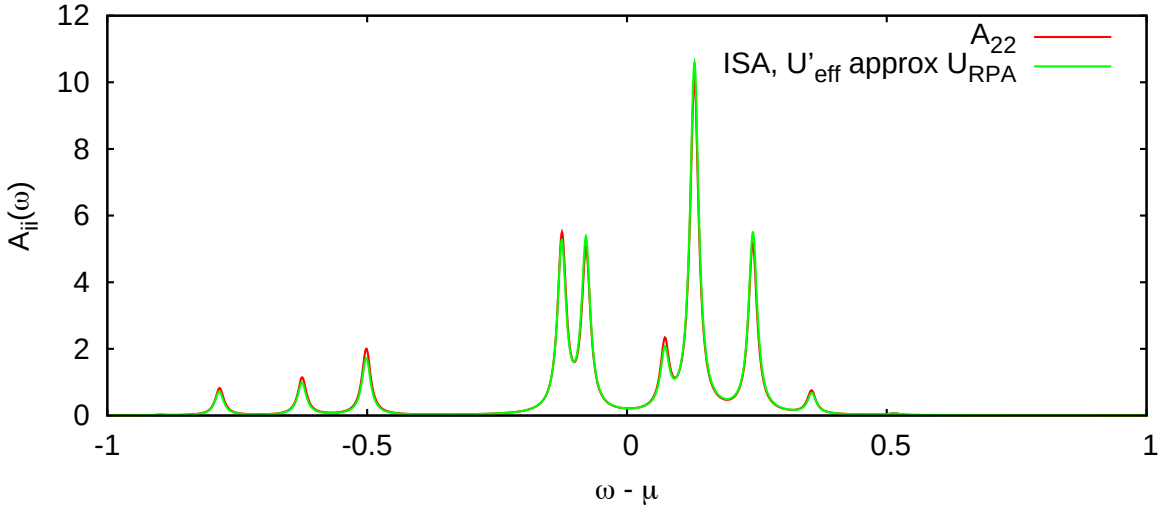
Figure 5.3.: Spectral function for the effective U obtained via RPA. Parameters: Set WrMcSd from Tab. 4.2, $\Delta = 1.0$. Full spectral function (red), effective single band within ISA, $U = U'_{\text{eff}}$. (dotted, green), c-RPA (dotted, green) and ISA with d-orbital occupation fixed at desired values, $U = U_{\text{eff}}$ (dotted, black).

In Fig. 5.3, we show the spectral function of the effective single-band system when U is taken from c-RPA. Here, much better agreement with the full spectrum is observed. The Mott gap is reproduced accurately, while the hopping is, as expected, slightly overestimated. The exact diagonalization for the r-electrons shows that the hopping is reduced by a factor of 0.999, so hopping reduction is negligible. In the following, we provide a detailed explanation of why the spectrum of the d-electrons is accurately reproduced by the effective single-band with U from c-RPA even though we are in a regime where d-electrons

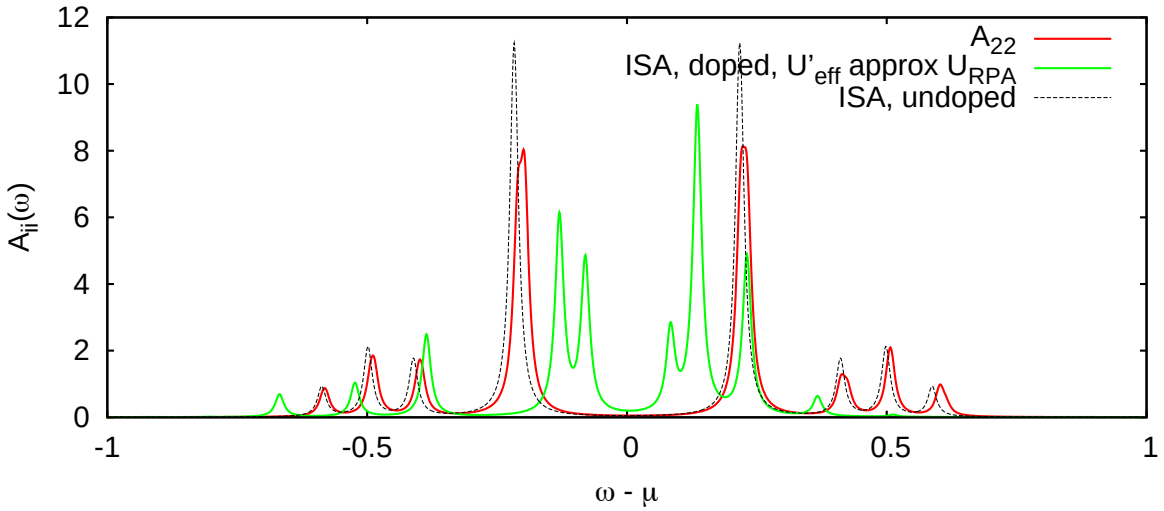
leave their band. First, we note that the spectral function of the d-electrons still looks like that of a single-band Hubbard system at half filling, although for the parameters used the system becomes unstable when two d-electrons occupy the same site. In other words, our effective single-band system with the U taken from c-RPA still accurately describes our system, although it does not correspond to a stable state. This is due to two effects. First, for the given parameters we calculate a probability for doubly occupied sites in the d-band of the full system of $D = 0.076$, which is quite small compared to the maximum of $D = 1/4$ for an uncorrelated system. We also find an overall occupancy of the d-band slightly less than 1/2, ($n_2 = 0.99935/2$). This means that the instability due to doubly occupied sites does not have a strong effect on the full system's ground state. Second, there is only a very small matrix element $t_{23} = t_{22}/1000$ between the d- and r-orbitals. Consequently, an electron added to the d-band in an inverse photoemission process cannot quickly leave this band. In other words, this energetically unfavorable state is meta-stable.

To confirm this explanation, we turn to a system that has been doped away from half filling. When we have one additional electron, the ground state of a single-band Hubbard model must contain at least one doubly occupied site. As we are in a parameter regime where this leads to an instability, we expect that in this case the resulting single-band spectrum will be a poor match to that of the full system. We will again introduce a small shift ε to one of the orbital energies to lift the degeneracy and average over the two spin types as in Sec. 4.2.4. We show the resulting spectra in Fig. 5.4, where we have one additional electron. For a parameter set where the doubly occupied d-orbital is stable, we find good agreement between the full system's spectral function and that of an effective single-band model (Fig. 5.4(a)). The Hubbard U was calculated within ISA, i.e., we use U'_{eff}. For the parameter set that showed the instability for doubly occupied d-orbitals, we show the resulting spectral function in Fig. 5.4(b). Here, U_{eff} was used. While the spectrum of the single-band system looks similar to that of the stable case (Fig. 5.4(a)), the spectrum of the full system's d-band looks more like that of the system at half filling, as can be seen from the dashed line in Fig. 5.4(b). For parameters where doubly occupied sites are unstable, the additional electron introduced by doping does not reside in the d- but is transferred to the r-band. Our calculations confirm that n_d remains at a value slightly below 1 for each site. There, it only serves to weakly renormalize the interaction by changing the energy of the r-electron ground state. Because of the instability of doubly occupied sites, the d-band remains at half filling.

To complete the study of this parameter set, we show the results for a system that has been doped *below* half filling in Fig. 5.5. Here, we expect that the destabilizing effect of doubly occupied sites plays a negligible role. Indeed, we first observe that the spectral function of the full system's d-band does not correspond to that of a system at half filling. Second, we see that the discrepancy that is due to using U'_{eff} instead of U_{eff} is less severe, because double occupations are very rare. It is worth mentioning at this point that in the single-band Hubbard model, there is a symmetry between particles and holes that leads to spectra for hole and electron systems to be the same. For our three-band system, however, we see that it can make a significant difference whether an electron is added to or removed from the system.



(a) Params. (WrMcSd) at $\Delta = 1.4$, stable



(b) Params. (WrMcSd) $\Delta = 1.0$, unstable

Figure 5.4.: Spectral function of a doped system (one additional electron) for parameter set (WrMcSd). Plot (a) shows the situation for $\Delta = 1.4$, where the doubly occupied d-orbital is stable (cf. Fig. 5.2(b)). Plot (b) shows the situation for $\Delta = 1.0$, where the doubly occupied d-orbital is unstable. The dashed line shows the spectral function of a system at half filling. As in the parameter set (WrMcSd) U_{eff} and $U_{\text{R}}A$ are very close, the spectra for RPA are omitted.

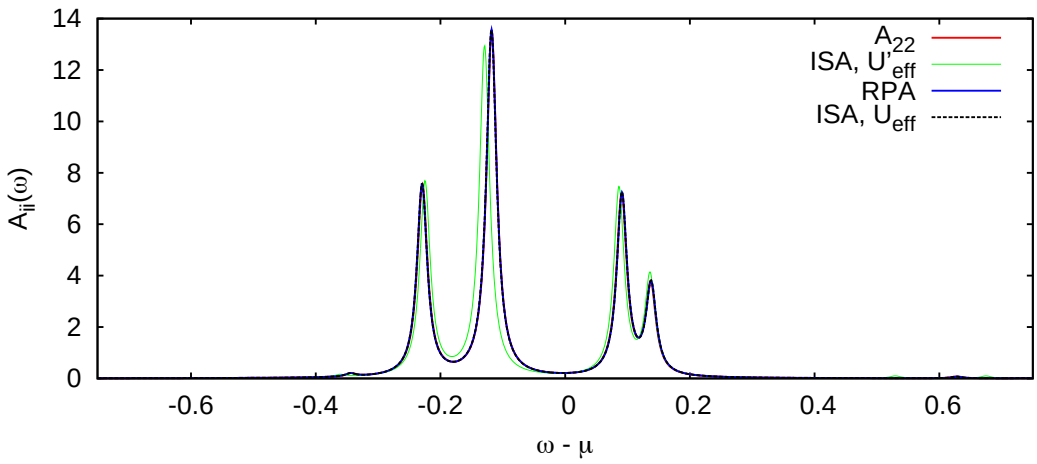


Figure 5.5.: Spectral function for a system with parameters (WrMcSd), $\Delta = 1$ doped away from half filling.

6. Summary

In this thesis, we have used the computational power of modern supercomputers to study renormalization effects that are modeled after those we expect in complicated realistic materials. With the exact ground state and excitation spectra at hand, we could quantitatively evaluate different approaches to obtaining effective Hubbard parameters. We devised a three-band model that is still exactly solvable with our computational methods yet is rich enough to show interesting screening effects. First, we demonstrated that by using a heuristic approach to screening, we can already accurately reproduce the Mott gap of the full system, but some deviations in the minor peaks remain. We then developed the theoretical background for the screening effects found in a Hubbard model due to external bands. Within our theoretical approach, we established that in addition to the well known screening of the Coulomb interaction, the hopping matrix elements have to be reduced as well. This hopping reduction then leads to an excellent description of the full spectrum using an effective single-band model. We have verified that the effective single-band description remains accurate for doped systems. Further studies of these interesting systems are worthwhile but were beyond the time frame of this work.

The hopping reduction resolves a question in the calculation of Hubbard parameters for realistic applications; often, the calculated effective U appeared too small to describe experimental observations. This is because the physics of the Hubbard model are mainly driven by the ratio U/t rather than by the precise values of U and t themselves. This ratio will be too small when the correct, reduced value for U is used together with the bare, unreduced value for t. The ratio can be corrected by either raising U or lowering t. We have established that the latter way is the correct one.

We have compared the ISA to another method of obtaining an effective Hubbard U, c-RPA. In general, we found good agreement between both approximations.

For the ISA, we still need to find solutions for the r-electrons. For realistic materials beyond our simple model, successful computational schemes, such as DFT, are available when the electrons are not too strongly correlated. It will be an interesting direction for future research to see how important the hopping reduction is in realistic materials.

A. Three Interacting Orbitals in Second Quantization

We start from the three orbitals 1, 2 and 3 as defined in 3.1. The hybridization t_{13} leads to new orbitals + and −. The old and new orbitals are related via

$$c_1^\dagger = \alpha c_-^\dagger + \beta c_+^\dagger \tag{A.1}$$
$$c_3^\dagger = \beta c_-^\dagger - \alpha c_+^\dagger \tag{A.2}$$

where $\alpha = \cos\frac{\theta}{2}$ and $\beta = \sin\frac{\theta}{2}$ with $\tan\theta = \frac{t_{13}}{\Delta}$. The destruction operators are obtained analogously. The new eigenenergies are $\pm\sqrt{\Delta^2 + t_{13}^2}$.

To rewrite the Hamiltonian (3.1) in the new basis, we have to rewrite the particle number operators n_i. While n_2 remains unchanged, the densities n_1 and n_3 are reformulated. We expect that this leads to exchange terms beyond simple density-density products. The densities themselves transform as follows.

$$\begin{aligned} n_1 &= \sum_\sigma n_{1\sigma} = \sum_\sigma c_{1\sigma}^\dagger c_{1\sigma} = \sum_\sigma \left[\alpha c_{-\sigma}^\dagger + \beta c_{+\sigma}^\dagger\right]\left[\alpha c_{-\sigma} + \beta c_{+\sigma}\right] \\ &= \sum_\sigma \left[\alpha^2 c_{-\sigma}^\dagger c_{-\sigma} + \beta^2 c_{+\sigma}^\dagger c_{+\sigma} + \alpha\beta\left(c_{-\sigma}^\dagger c_{+\sigma} + c_{+\sigma}^\dagger c_{-\sigma}\right)\right] \\ &= \alpha^2 n_- + \beta^2 n_+ + \alpha\beta\sum_\sigma \left(c_{-\sigma}^\dagger c_{+\sigma} + c_{+\sigma}^\dagger c_{-\sigma}\right) \end{aligned} \tag{A.3}$$

$$n_3 = \beta^2 n_- + \alpha^2 n_+ - \alpha\beta\sum_\sigma \left(c_{-\sigma}^\dagger c_{+\sigma} + c_{+\sigma}^\dagger c_{-\sigma}.\right) \tag{A.4}$$

Note that the particle number operators, rewritten in the new basis, contain hopping terms between the orbitals + and −. This leads to interesting effects when interaction terms – involving products of particle number operators – are calculated.

A.1. Intra-orbital Interactions

The on-site interaction involves products of the form $n_{1\uparrow}n_{1\downarrow}$, which for Fermions can be rewritten as $\frac{1}{2}n_1(n_1-1)$. We therefore look at that product. From (A.3) we obtain

$$\begin{aligned} n_1(n_1-1) = & \\ & \left(\alpha^2 n_- + \beta^2 n_+ + \alpha\beta \sum_\sigma c^\dagger_{-\sigma}c_{+\sigma} + c^\dagger_{+\sigma}c_{-\sigma}\right) \\ \otimes & \left(\alpha^2 n_- + \beta^2 n_+ + \alpha\beta \sum_\sigma c^\dagger_{-\sigma}c_{+\sigma} + c^\dagger_{+\sigma}c_{-\sigma} - 1\right) \\ & = \alpha^4 n_-^2 + \beta^4 n_+^4 + 2\alpha^2\beta^2 n_- n_+ - \alpha^2 n_- - \beta^2 n_+ + \text{ hopping terms.} \end{aligned} \tag{A.5}$$

The terms that are explicitly given in A.5 are direct electron-electron interaction terms between electrons in the $+$ and $-$ orbitals and result from multiplying the density terms in (A.3). To calculate the hopping terms, which result from products involving the hopping part of (A.3), some operator algebra is necessary. We approach the summands one by one and omit α and β for the moment.

$$
\begin{aligned}
T_1 &= n_- \sum_\sigma \left(c^\dagger_{-\sigma} c_{+\sigma} + c^\dagger_{+\sigma} c_{-\sigma} \right) = \sum_{\sigma,\sigma'} c^\dagger_{-\sigma'} c_{-\sigma'} \left(c^\dagger_{-\sigma} c_{+\sigma} + c^\dagger_{+\sigma} c_{-\sigma} \right) \\
&= \sum_\sigma c^\dagger_{-\sigma} c_{-\sigma} \left(c^\dagger_{-\sigma} c_{+\sigma} + c^\dagger_{+\sigma} c_{-\sigma} \right) + \sum_\sigma c^\dagger_{-\bar{\sigma}} c_{-\bar{\sigma}} \left(c^\dagger_{-\sigma} c_{+\sigma} + c^\dagger_{+\sigma} c_{-\sigma} \right) \\
&= \sum_\sigma c^\dagger_{-\sigma} c_{+\sigma} + \sum_\sigma n_{-\bar{\sigma}} \left(c^\dagger_{-\sigma} c_{+\sigma} + c^\dagger_{+\sigma} c_{-\sigma} \right) \qquad \text{(A.6)} \\
T_2 &= \left[\sum_\sigma \left(c^\dagger_{-\sigma} c_{+\sigma} + c^\dagger_{+\sigma} c_{-\sigma} \right) \right] n_- \\
&= \sum_\sigma c^\dagger_{+\sigma} c_{-\sigma} + \sum_\sigma n_{-\bar{\sigma}} \left(c^\dagger_{-\sigma} c_{+\sigma} + c^\dagger_{+\sigma} c_{-\sigma} \right) = T_1^\dagger \qquad \text{(A.7)} \\
T_3 &= n_+ \sum_\sigma \left(c^\dagger_{-\sigma} c_{+\sigma} + c^\dagger_{+\sigma} c_{-\sigma} \right) \\
&= \sum_\sigma c^\dagger_{+\sigma} c_{+\sigma} \left(c^\dagger_{-\sigma} c_{+\sigma} + c^\dagger_{+\sigma} c_{-\sigma} \right) + \sum_\sigma c^\dagger_{+\bar{\sigma}} c_{+\bar{\sigma}} \left(c^\dagger_{-\sigma} c_{+\sigma} + c^\dagger_{+\sigma} c_{-\sigma} \right) \\
&= \sum_\sigma c^\dagger_{+\sigma} c_{-\sigma} + \sum_\sigma n_{+\bar{\sigma}} \left(c^\dagger_{-\sigma} c_{+\sigma} + c^\dagger_{+\sigma} c_{-\sigma} \right) \qquad \text{(A.8)} \\
T_4 &= \sum_\sigma c^\dagger_{-\sigma} c_{+\sigma} + \sum_\sigma n_{+\bar{\sigma}} \left(c^\dagger_{-\sigma} c \, {+\sigma} + c^\dagger_{+\sigma} c_{-\sigma} \right) \qquad \text{(A.9)} \\
T_5 &= \left[\sum_\sigma c^\dagger_{-\sigma} c_{+\sigma} + c^\dagger_{+\sigma} c_{-\sigma} \right]^2 = \sum_{\sigma\sigma'} \left(c^\dagger_{-\sigma} c_{+\sigma} + c^\dagger_{+\sigma} c_{-\sigma} \right) \left(c^\dagger_{-\sigma'} c_{+\sigma'} + c^\dagger_{+\sigma'} c_{-\sigma'} \right) \\
&= \sum_\sigma \left[n_{-\sigma} + n_{+\sigma} - 2 n_{-\sigma} n_{+\sigma} + c^\dagger_{-\bar{\sigma}} c^\dagger_{-\sigma} c_{+\sigma} c_{+\bar{\sigma}} + c^\dagger_{+\bar{\sigma}} c^\dagger_{+\sigma} c_{-\sigma} c_{-\bar{\sigma}} + 2 c^\dagger_{-\bar{\sigma}} c^\dagger_{+\sigma} c_{-\sigma} c_{=\bar{\sigma}} \right] \\
& \qquad \text{(A.10)} \\
T_6 &= \sum_\sigma \left(c^\dagger_{+\sigma} c_{-\sigma} + c^\dagger_{-\sigma} c_{+\sigma} \right) \qquad \text{(A.11)}
\end{aligned}
$$

The final hopping term reads

$$
\alpha^2 \alpha\beta (T_1 + T_2) + \beta^2 \alpha\beta (T_3 + T_4) + \alpha^2 \beta^2 T_5 - T_6. \qquad \text{(A.12)}
$$

Together with the other terms in (A.5), and noting that $\alpha^2 + \beta^2 = 1$ we can simplify the final result a bit. Let us go through it in a systematic way.

- The first-order hopping term cancels out in T_1, T_2, T_3, T_4 and T_6.
- The terms only containing powers of n_- yield

$$
\alpha^4 n_-^2 - \alpha^2 n_- + \alpha^2 \beta^2 n_- = \alpha^4 n_-^2 - \alpha^2 n_- + \alpha^2 (1 - \alpha^2) n_- = \alpha^4 n_- (n_- - 1) \qquad \text{(A.13)}
$$

- Analogously, the terms only containing powers of n_+ yield

$$\beta^4 n_+(n_+ - 1) \tag{A.14}$$

- Because of the sum in T_5 we get an interesting form for the interaction between the two orbitals.

$$2\alpha^2\beta^2 \left(n_- n_+ - \sum_\sigma n_{-\sigma} n_{+\sigma} \right) = 2\alpha^2\beta^2 (n_{-\uparrow} n_{+\downarrow} + n_{-\downarrow} n_{+\uparrow}) \tag{A.15}$$

- The next term describes Hopping of one electron augmented by the presence of another electron. There is, however, not much to simplify.

$$2\alpha\beta \sum_\sigma \left(\alpha^2 n_{-\bar{\sigma}} + \beta^2 n_{+\bar{\sigma}} \right) \left(c^\dagger_{+\sigma} c_{-\sigma} + c^\dagger_{-\sigma} c_{+\sigma} \right) \tag{A.16}$$

- Finally, there is a hopping term in T_5 that describes simultaneous hopping of two electrons between orbitals + and −.

$$\sum_\sigma \left[c^\dagger_{-\bar{\sigma}} c^\dagger_{-\sigma} c_{+\sigma} c_{+\bar{\sigma}} + c^\dagger_{+\bar{\sigma}} c^\dagger_{+\sigma} c_{-\sigma} c_{-\bar{\sigma}} + 2 c^\dagger_{-\bar{\sigma}} c^\dagger_{+\sigma} c_{-\sigma} c_{+\bar{\sigma}} \right] \tag{A.17}$$

The result for the term $n_3(n_3 - 1)$ is readily obtained by exchanging α and β in the previous derivations.

A.2. Inter-orbital Interaction

So far, we looked at the interaction between the orbitals + and −. We now turn to their interaction with the central band. In contrast to the previous results, the terms do not become very complex because the Fermi operators for orbitals + and − anti-commute with the ones for the central orbital.

$$\begin{aligned} n_2 n_1 &= n_2 \left[\alpha^2 n_- + \beta^2 n_+ + \alpha\beta \sum_\sigma \left(c^\dagger_{-\sigma} c_{+\sigma} + c^\dagger_{+\sigma} c_{-\sigma} \right) \right] \\ &= \alpha^2 n_- n_2 + \beta^2 n_2 n_+ + n_2 \alpha\beta \sum_\sigma \left(c^\dagger_{-\sigma} c_{+\sigma} + c^\dagger_{+\sigma} c_{-\sigma} \right) \end{aligned} \tag{A.18}$$

In the same way, we obtain the expression for $n_2 n_3$. The first two terms describe the interaction between the central and the outer orbitals. The last term describes hopping between the outer bands that depends on the occupation of the central orbital. This is crucial for the screening effect we have in mind for our model as it describes how the presence of electrons in the central orbital on a site facilitates hopping from the lower to the upper band.

We want to emphasize that for all practical purposes we prefer the description in terms of the old orbitals 1 and 3. With these orbitals, the interaction part of the many-body Hamiltonian is diagonal whereas the orbitals + and − introduce complicated off-diagonal matrix elements. They are important from a conceptually point of view as they demonstrate how the hybridization t_{13} introduces complex interactions not only between the outer bands but also with the central band.

B. Exact Solution of Atomic Limit

In the atomic limit, we only have to consider a single site. The number of d-electrons is fixed, so we have to specify only the state of the r-electrons. For the many-body basis, we choose $|\uparrow\downarrow,\cdot\rangle$, $|\cdot,\uparrow\downarrow\rangle$, $|\uparrow,\downarrow\rangle$ and $|\downarrow,\uparrow\rangle$. The resulting Hamiltonian is a 4×4 matrix. We concentrate on a simple case where all U_{ij} are zero except for U_{12}, which yields

$$H = \begin{pmatrix} -2\Delta + 2U_{12}n_2 & 0 & t & t \\ 0 & 2\Delta & t & t \\ t & t & U_{12}n_2 & 0 \\ t & t & 0 & U_{12}n_2 \end{pmatrix} \tag{B.1}$$

where we write t instead of t_{13}. We can incorporate a non-zero U_{23} as well by replacing U_{12} with $U_{12} - U_{23}$ whenever it occurs, but we will refrain from doing so for the sake of simplicity. By shifting the zero energy level down by $U_{12}n_2$, we obtain

$$H = \begin{pmatrix} -A & 0 & t & t \\ 0 & A & t & t \\ t & t & 0 & 0 \\ t & t & 0 & 0 \end{pmatrix} + U_{12}n_2 \cdot \mathbb{1} \tag{B.2}$$

with $A = 2\Delta - U_{12}n_2$. The matrix can be diagonalized using a computer algebra program such as MATHEMATICA. The resulting eigenenergies are

$$\begin{aligned} E_0 &= -\sqrt{A^2+4t^2} & E_1 &= 0 \\ E_2 &= 0 & E_3 &= \sqrt{A^2+4t^2} \end{aligned}, \tag{B.3}$$

and the resulting normalized eigenvectors are

$$\begin{aligned} \mathbf{v}_0 &= \frac{1}{\sqrt{A^2+4t^2}} \begin{pmatrix} -\frac{1}{2}\sqrt{A^2+4t^2} - \frac{1}{2}A \\ -\frac{1}{2}\sqrt{A^2+4t^2} + \frac{1}{2}A \\ t \\ t \end{pmatrix}, \quad \mathbf{v}_3 = \frac{1}{\sqrt{A^2+4t^2}} \begin{pmatrix} \frac{1}{2}\sqrt{A^2+4t^2} - \frac{1}{2}A \\ \frac{1}{2}\sqrt{A^2+4t^2} + \frac{1}{2}A \\ t \\ t \end{pmatrix} \\ \mathbf{v}_1 &= \frac{1}{\sqrt{A^2+2t^2}} \begin{pmatrix} t \\ -t \\ 0 \\ A \end{pmatrix}, \qquad \mathbf{v}_2 = \frac{1}{\sqrt{A^2+2t^2}} \begin{pmatrix} t \\ -t \\ A \\ 0 \end{pmatrix}, \end{aligned} \tag{B.4}$$

To obtain the occupation of the lower orbital in the ground state, we use the ground state vector to compute the expectation value of n_1, whose matrix form in the chosen basis is

$\mathrm{diag}(2,0,1,1)$. The result is

$$n_1 = 1 + \frac{A}{\sqrt{A^2+4t^2}}. \tag{B.5}$$

Note that A also depends on n_2.

B.1. Discontinuities

Discontinuities in the occupancy of orbitals is due to parameters that lead to d-electrons leaving their orbital. In this section, we will give an example of how the boundary region for such a discontinuity can be computed. We consider the case where $n_2 = 1$. From (B.3) we now the ground state energy to be

$$E_0 = -\sqrt{A^2+4t^2} + U_{12}n_2, \tag{B.6}$$

with $A = -2\Delta - 2U_{12}$. Now we consider the case where the d-electron leaves its orbital. Then, the only possible basis states for the r-electrons are $|\uparrow\downarrow,\uparrow\rangle$ and $|\uparrow,\uparrow\downarrow\rangle$, The Hamiltonian is

$$\begin{pmatrix} -\Delta & t \\ t & \Delta \end{pmatrix}, \tag{B.7}$$

whose ground state energy is

$$E_0' = -\sqrt{\Delta^2+t^2} \tag{B.8}$$

The d-electron leaves its orbital once $E_0' < E_0$. Thus, we can compute the position of the discontinuity by setting $E_0' = E_0$ and solving for U_{12} as a function of t. This can be done by hand and yields

$$U_{12}^{\mathrm{crit}} = \frac{3}{2}\frac{\Delta^2+t^2}{\sqrt{\Delta^2+t^2}+2\Delta}. \tag{B.9}$$

For the discontinuity that arises when one of two d-electrons leaves its orbital, the r-electron states are the same, but because of the presence of another d-electron, the Hamiltonian reads

$$H = \begin{pmatrix} -\Delta + 2U_{12}n_2 & t \\ t & \Delta + U_{12}n_2 \end{pmatrix}, \tag{B.10}$$

whose eigenvalues are

$$\pm\sqrt{\left(\Delta - \frac{1}{2}U_{12}\right)^2 + t^2} + \frac{3}{2}U_{12} \tag{B.11}$$

For the transition from 2 d-electrons to 1 d-electron, we had to solve a biquadratic equation. Although MATHEMATICA returns a closed formula, it is too complicated to give any insight. We use the program to expand the solution to second order in t/Δ and insert concrete values for Δ and U_{22}. For $\Delta = 1.4$ and $U_{22} = 0.6$, we obtain

$$U_{12}^{\mathrm{crit}} = 0.4 + 0.2917t^2. \tag{B.12}$$

For the discontinuity that arises when the second d-electron leaves its orbital, we can obtain a closed formula again. The resulting Hilbert space contains only the state $|\uparrow\downarrow,\uparrow\downarrow\rangle$, whose energy is 0. By equating this with (B.11), we obtain

$$U_{12}^{\text{crit}} = -\frac{\Delta}{4} + \sqrt{\frac{9}{16}\Delta^2 + \frac{1}{2}t^2}. \tag{B.13}$$

B.2. Change in Density

We are interested in computing $\delta n_1 = n_1(n_2 = 1) - n_1(n_2 = 2)$. We already have a closed formula (B.5) for n_1, so we have

$$\delta n_1 = \frac{2\Delta - U_{12}}{\sqrt{(2\Delta - U_{12})^2 + 4t^2}} - \frac{2\Delta - 2U_{12}}{\sqrt{(2\Delta - 2U_{12})^2 + 4t^2}} - \tag{B.14}$$

Unfortunately, we cannot simplify this expression. To gain some insight into its behavior, we expand it in a power series for small U_{12}/Δ and t/Δ. By writing U and t in units of Δ, we obtain

$$\frac{\delta n_1}{\Delta} \approx \left(\frac{t_{13}}{\Delta}\right)^2 \left(\frac{U_{12}}{2\Delta} + \frac{9}{8}\left(\frac{U_{12}}{\Delta}\right)^2\right). \tag{B.15}$$

B.3. Screening

We can use our expression for the ground state energy to obtain an expression for the screening of the Hubbard U. Let A_i be the expression A for $n_2 = i$. Then, clearly,

$$\begin{aligned} U_{22} - U &= U_{22} - \mathcal{E}_0(2) + 2\mathcal{E}_0(1) - mathcal E_0(0) \\ &= -\sqrt{A_2^2 + 4t^2} + 2\sqrt{A_1^2 + 4t^2} - \sqrt{A_0^2 + 4t^2} \\ &= -\sqrt{(2\Delta - 2U_{12})^2 + 4t^2} + 2\sqrt{(2\Delta - U_{12})^2 + 4t^2} - 2\sqrt{\Delta^2 + t^2} \\ &= \Delta\left[-\sqrt{(2 - 2\tilde{U})^2 + 4\tilde{t}^2} + 2\sqrt{(2 - \tilde{U})^2 + 4\tilde{t}^2} - 2\sqrt{1 + \tilde{t}^2}\right] \end{aligned} \tag{B.16}$$

where $\tilde{U} = U/\Delta$ and $\tilde{t} = t/\Delta$. A Taylor expansion of the square bracket in $\tilde{U}$ and $\tilde{t}$ yields

$$U_{22} - U \approx \Delta\tilde{t}^2\tilde{U}^2 = \frac{t^2 U_{12}^2}{\Delta^3} + \mathcal{O}\left(\frac{1}{\Delta^4}\right). \tag{B.17}$$

C. Antisymmetrization

Let $c^\dagger_{d,n}$ and $c^\dagger_{r,m}$ be a complete set of creation operators for the d- and r-electrons, respectively. We can use these to create both completely and partially antisymmetrized states. Let

$$\hat{\psi}^\dagger_d = \sum_n a_{n,\sigma} \prod_{\sigma,n} \left(c^\dagger_{n,\sigma}\right)^{n_{n,\sigma}} \tag{C.1}$$

be the operator that creates $|\psi_d\rangle$ and define $\hat{\psi}^\dagger_r$ analogously. Then the completely antisymmetrized combination of these states is

$$|\Psi\rangle_- = \hat{\psi}^\dagger_d \hat{\psi}^\dagger_r |0\rangle \tag{C.2}$$

and the partially antisymmetrized combination is

$$|\Psi\rangle_\otimes = \left[\hat{\psi}^\dagger_d |0\rangle\right] \otimes \left[\hat{\psi}^\dagger_r |0\rangle\right] = |\psi_d\rangle \otimes |\psi_r\rangle \tag{C.3}$$

We will show that matrix elements of operators that respect the separation of the d- and the r-electrons do not depend on the choice of the combination. Let $\hat{A}$ be an arbitrary operator with the only restriction that it preserves N_d and N_r. We can expand $\hat{A}$ in a series of products of creation and annihilation operators. Using the linearity of matrix elements, we can without loss of generality restrict $\hat{A}$ to product form. Because the particle operators anti-commute for different electron types, we can further demand that $\hat{A}$ is of the form $\hat{A} = \hat{A}_d \hat{A}_r$ where $\hat{A}_d$ contains only particle operators of the d-electrons and $\hat{A}_r$ contains only those for r-electrons.

We begin to compute the matrix elements of $\hat{A}$ for the partially antisymmetrized case.

$$\left(\langle\psi'_d| \otimes \bar{\psi'_r}\right) \hat{A}_d \hat{A}_r \left(|\psi_r\rangle \otimes |\psi_d\rangle\right) = \langle\psi'_d| \hat{A}_d |\psi_d\rangle \langle\psi'_r| \hat{A}_r |\psi_r\rangle \tag{C.4}$$

Here, we make use of the fact that $\hat{A}_d$ does not act on the r-electrons and $\hat{A}_r$ does not act on the d-electrons. The completely antisymmetrized case is slightly more involved.

$$\begin{aligned}
\langle 0|\, \psi'_d\, \psi'_r\, \hat{A}_d \hat{A}_r\, \psi^\dagger_r \psi^\dagger_d\, |0\rangle &= \langle 0|\, \psi'_d\, \hat{A}_d\, \psi^\dagger_d \psi'_r\, \hat{A}_r\, \psi^\dagger_r\, |0\rangle && (*) \\
&= \langle 0|\, \psi'_d\, \hat{A}_d\, \psi^\dagger_d \sum_n |\Psi_n\rangle \langle\Psi_n|\, \psi'_r\, \hat{A}_r\, \psi^\dagger_r\, |0\rangle && (**) \\
&= \langle 0|\, \psi'_d\, \hat{A}_d\, \psi^\dagger_d\, |0\rangle \langle 0|\, \psi'_r\, \hat{A}_r\, \psi^\dagger_r\, |0\rangle && (***) \\
&= \langle\psi'_d| \hat{A}_d |\psi_d\rangle \langle\psi'_r| \hat{A}_r |\psi_r\rangle && (C.5)
\end{aligned}$$

In (*) we used that $\hat{A}_d$ and $\hat{A}_r$ consist of an even number of particle operators and that the particle operators for different types anti-commute. In (**) we inserted a complete

orthonormal basis for the entire Fock space, that is $\mathcal{F} = \sum_N \mathcal{H}^N$, where we demand that each basis state $|\Psi_n\rangle$ has a definite number of electrons. To get from (**) to (***), we note that the product of operators $\psi'_d \hat{A}_d \psi^\dagger_d$ does not change the number of electrons in a state and that the overlap of states with different, but definite, particle numbers is zero. Therefore, the only state $|\Psi_n\rangle$ that does not lead to zero overlap with $\langle 0|$ after applying the operators to it is $|0\rangle$. Finally, we are left with the same result for the matrix element as in (C.4)

D. Various Definitions

This chapter contains various formal definitions.

D.1. Many-particle Operators

We want to define many-particle creation and annihilation operators that act like

$$\hat{\psi}_0^\dagger(\mathbf{n_d})\,|0\rangle = |\psi_0(\mathbf{n_d})\rangle \tag{D.1}$$

$$\hat{\psi}_0(\mathbf{n_d})\,|\psi_0(\mathbf{n_d})\rangle = |0\rangle \tag{D.2}$$

Generally, a state $|\psi_0(\mathbf{n_d})\rangle$ can be expanded in the configuration basis as

$$\begin{aligned}|\psi_0(\mathbf{n_d})\rangle &= \sum_{\mathbf{n_r}} \psi_0(\mathbf{n_r};\mathbf{n_d})\,|\mathbf{n_r}\rangle \\ &= \sum_{\mathbf{n_r}} \psi_0(\mathbf{n_r};\mathbf{n_d}) \prod_{\nu=1}^{N} \left(c_\nu^\dagger\right)^{n_\nu} |0\rangle \\ &=: \hat{\psi}_0^\dagger(\mathbf{n_d})\,|0\rangle \end{aligned} \tag{D.3}$$

where ν is a generalized index enumerating the states of the r-electrons. From this, we readily obtain the operator:

$$\hat{\psi}_0^\dagger(\mathbf{n_d}) = \sum_{\mathbf{n_r}} \psi_0(\mathbf{n_r};\mathbf{n_d}) \prod_{\nu=1}^{N} \left(c_\nu^\dagger\right)^{n_\nu}. \tag{D.4}$$

The destruction operator is the adjoint of the creation operator:

$$\hat{\psi}_0(\mathbf{n_d}) = \sum_{\mathbf{n_r}} \psi_0(\mathbf{n_r};\mathbf{n_d}) \prod_{\nu=N}^{1} \left(c_\nu\right)^{n_\nu}. \tag{D.5}$$

Note that the ordering of the particle operators is reversed and that we assume the coefficients $\psi_0(\mathbf{n_r};\mathbf{n_d})$ to be real valued. We can verify that the definition for $\hat{\psi}_0(\mathbf{n_d})$ is correct by applying it to $|\psi_0(\mathbf{n_d})\rangle$:

$$\begin{aligned}\hat{\psi}_0(\mathbf{n_d})\,|\psi_0(\mathbf{n_d})\rangle &= \hat{\psi}_0(\mathbf{n_d})\hat{\psi}_0^\dagger(\mathbf{n_d})\,|0\rangle \\ &= \sum_{\mathbf{n_r},\mathbf{n_r'}} \psi_0(\mathbf{n_r};\mathbf{n_d})\psi_0(\mathbf{n_r'};\mathbf{n_d}) \prod_{\nu=N}^{1} \left(c_\nu\right)^{n_\nu} \prod_{\mu=1}^{N} \left(c_\mu^\dagger\right)^{n_\mu'} |0\rangle\,. \end{aligned}$$

The total number of creation and annihilation operators is equal in both products; only when each creation operator is matched by the corresponding annihilation operator will the result of them acting on $|0\rangle$ be non-zero.

$$= \sum_{\mathbf{n_r}} \psi_0(\mathbf{n_r};\mathbf{n_d})^2 \prod_{\nu=N}^{1} \left(c_\nu\right)^{n_\nu} \prod_{\mu=1}^{N} \left(c_\mu^\dagger\right)^{n_\mu} |0\rangle \tag{D.6}$$

Because of the reverse ordering of the first product, the operators c_1 and $c_1^\dagger$ occur as a pair and can be moved to the end of the overall product without introducing a Fermi sign. After this step, there is a pair $c_2 c_2^\dagger$ which can be moved to the end. By iterating this process for all N pairs, we have

$$\begin{aligned} &= \sum_{\mathbf{n_r}} \psi_0(\mathbf{n_r};\mathbf{n_d})^2 \prod_{\nu=1}^{N} \left(c_\nu c_\nu^\dagger\right) |0\rangle = \sum_{\mathbf{n_r}} \psi_0(\mathbf{n_r};\mathbf{n_d})^2 \prod_{\nu=1}^{N} \left(1 - c_\nu^\dagger c_\nu\right) |0\rangle \\ &= \sum_{\mathbf{n_r}} \psi_0(\mathbf{n_r};\mathbf{n_d})^2 |0\rangle = |0\rangle \,. \end{aligned} \tag{D.7}$$

In the last step we used that the $|\psi_0(\mathbf{n_d})\rangle$ is normalized. We see that $\hat{\psi}_0(\mathbf{n_d})$ behaves as desired.

The intended use for the many-particle operators defined in this section is to create states $|\psi_0(\mathbf{n_d})\rangle$ by

$$\hat{\psi}_0^\dagger(\mathbf{n_d}) |0\rangle \tag{D.8}$$

and to switch from a state $|\psi_0(\mathbf{n_d})\rangle$ to a state $|\psi_0(\mathbf{n'_d})\rangle$ by

$$\hat{\psi}_0^\dagger(\mathbf{n'_d})\psi_0(\mathbf{n_d}) |\psi_0(\mathbf{n_d})\rangle \,. \tag{D.9}$$

Applying the many-particle operators in any other way leads to complicated results that shall not concern us here.

D.2. Instantaneous Coupling Operators

We define

$$\hat{\Psi}_-(i\sigma) = \sum_{\substack{\mathbf{n_d} \\ n_{di\sigma}=1}} \hat{\psi}_0^\dagger(\mathbf{n_d^{-i\sigma}})\hat{\psi}_0(\mathbf{n_d})\hat{P}_d(\mathbf{n_d}) \tag{D.10}$$

as the operator that switches the r-electron configuration from $|\psi_0(\mathbf{n_d})\rangle$ to $|\psi_0(\mathbf{n_d^{-i\sigma}})\rangle$ where $\mathbf{n_d^{-i\sigma}}$ is the configuration that arises from $\mathbf{n_d}$ when the electron in state $i\sigma$ is removed. In the same way, we define

$$\hat{\Psi}_+(i\sigma) = \sum_{\substack{\mathbf{n_d} \\ n_{di\sigma}=0}} \hat{\psi}_0^\dagger(\mathbf{n_d^{+i\sigma}})\hat{\psi}_0(\mathbf{n_d})\hat{P}_d(\mathbf{n_d}) \tag{D.11}$$

as the operator for the case where an electron in state $i\sigma$ is added. One expects that $\hat{\Psi}_-^\dagger(i\sigma) = \hat{\Psi}_+(i\sigma)$. This is not hard to show:

$$\hat{\Psi}_-^\dagger(i\sigma) = \sum_{\substack{\mathbf{n_d} \\ n_{di\sigma}=1}} \hat{P}_d(\mathbf{n_d})\hat{\psi}_0^\dagger(\mathbf{n_d})\hat{\psi}_0(\mathbf{n_d^{-i\sigma}})$$

We substitute the sum index from $\mathbf{n_d}$ to $\mathbf{n_d^{-i\sigma}}$

$$= \sum_{\substack{\mathbf{n_d} \\ n_{di\sigma}=0}} \hat{P}_d(\mathbf{n_d^{+i\sigma}})\hat{\psi}_0^\dagger(\mathbf{n_d^{+i\sigma}})\hat{\psi}_0(\mathbf{n_d})$$

Next, we want to move the projector to the right side. It singles out the states corresponding to $\mathbf{n_d^{+i\sigma}}$ *after* the change in configuration has taken place. This is equivalent to singling out $\mathbf{n_d}$ *before* the configuration is changed, which leaves us with

$$= \sum_{\substack{\mathbf{n_d} \\ n_{di\sigma}=0}} \hat{\psi}_0^\dagger(\mathbf{n_d^{+i\sigma}})\hat{\psi}_0(\mathbf{n_d})\hat{P}_d(\mathbf{n_d}) = \hat{\Psi}_+(i\sigma). \tag{D.12}$$

D.2.1. Commutation Relations

Let us briefly discuss the commutation relations of the modified operators. We begin with the anti-commutator of two destruction operators, $\{\tilde{c}_{i\sigma}\tilde{c}_{j\sigma'}\}$, and study its effect on the basis states of $\mathcal{H}_{\mathrm{ISA}}$. For $i\sigma = j\sigma'$, we already know that the result will be 0; we therefore focus on the case where $i\sigma \neq j\sigma'$.

$$\begin{aligned}
\{\tilde{c}_{i\sigma}\tilde{c}_{j\sigma'}\}\,|\mathbf{n_d}\rangle\,|\psi_0(\mathbf{n_d})\rangle = \;& c_{i\sigma}\Psi_-(i\sigma)c_{j\sigma'}\Psi_-(j\sigma')\,|\mathbf{n_d}\rangle\,|\psi_0(\mathbf{n_d})\rangle \\
& + c_{j\sigma'}\Psi_-(j\sigma')c_{i\sigma}\Psi_-(i\sigma)\,|\mathbf{n_d}\rangle\,|\psi_0(\mathbf{n_d})\rangle \\
= \;& c_{i\sigma}\Psi_-(i\sigma)(-1)^{\pi(j\sigma';\mathbf{n_d})}n_{d,j,\sigma'}\,|\mathbf{n^{-j\sigma'}}\rangle\,|\psi_0(\mathbf{n^{-j\sigma'}}\rangle \\
& + c_{j\sigma'}\Psi_-(j\sigma')(-1)^{\pi(i\sigma;\mathbf{n_d})}n_{d,i,\sigma}\,|\mathbf{n^{-i\sigma}}\rangle\,|\psi_0(\mathbf{n^{-i\sigma}}\rangle \\
= \;& (-1)^{\pi(i\sigma;\mathbf{n_d^{-j\sigma'}})}n_{d,i,\sigma}(-1)^{\pi(j\sigma';\mathbf{n_d})}n_{d,j,\sigma'}\,|\mathbf{n_d^{-i\sigma,-j\sigma'}}\rangle\,|\psi_0(\mathbf{n_d^{-i\sigma,-j\sigma'}})\rangle \\
& + (-1)^{\pi(j\sigma';\mathbf{n_d^{-i\sigma}})}n_{d,j,\sigma}(-1)^{\pi(i\sigma;\mathbf{n_d})}n_{d,i,\sigma}\,|\mathbf{n_d^{-i\sigma,-j\sigma'}}\rangle\,|\psi_0(\mathbf{n_d^{-i\sigma,-j\sigma'}})\rangle
\end{aligned} \tag{D.13}$$

The two summands are equal except for the Fermi sign. Let, without loss of generality, be $i\sigma$ the state that comes in our ordering of states before $j\sigma'$. Then

$$\begin{aligned}
\pi(i\sigma;\mathbf{n_d}) &= \pi(i\sigma;\mathbf{n_d^{-j\sigma'}}) \\
\pi(j\sigma';\mathbf{n_d}) &= 1 + \pi(i\sigma;\mathbf{n_d^{-i\sigma}})
\end{aligned} \tag{D.14}$$

that is, the number of commutations needed to get an operator for $i\sigma$ in place does not depend on whether or not $j\sigma'$ is occupied, but for $j\sigma'$, we need one more commutation when $i\sigma$ is occupied. With this, it immediately follows that

$$\{\tilde{c}_{i\sigma}, \tilde{c}_{j\sigma'}\} = 0$$

and, with an analogous derivation,

$$\{\tilde{c}^\dagger_{i\sigma}, \tilde{c}^\dagger_{j\sigma'}\} = 0. \tag{D.15}$$

Let us now turn to the mixed anticommutator, $\{\tilde{c}_{i\sigma}, \tilde{c}^\dagger_{j\sigma'}\}$, and begin with the case $i\sigma \neq j\sigma'$. Then, we have

$$\begin{aligned}
\{\tilde{c}_{i\sigma}\tilde{c}^\dagger_{j\sigma'}\}\,|\mathbf{n_d}\rangle\,|\psi_0(\mathbf{n_d})\rangle = \quad & c_{i\sigma}\Psi_-(i\sigma)c^\dagger_{j\sigma'}\Psi_+(j\sigma')\,|\mathbf{n_d}\rangle\,|\psi_0(\mathbf{n_d})\rangle \\
& + c^\dagger_{j\sigma'}\Psi_+(j\sigma')c_{i\sigma}\Psi_-(i\sigma)\,|\mathbf{n_d}\rangle\,|\psi_0(\mathbf{n_d})\rangle \\
= \quad & c_{i\sigma}\Psi_-(i\sigma)(-1)^{\pi(j\sigma';\mathbf{n_d})}(1-n_{d,j,\sigma'})\,|\mathbf{n_d^{+j\sigma'}}\rangle\,|\psi_0(\mathbf{n_d^{+j\sigma'}})\rangle \qquad (*) \\
& + c^\dagger_{j\sigma'}\Psi_+(j\sigma')(-1)^{\pi(i\sigma;\mathbf{n_d})}n_{d,i,\sigma}\,|\mathbf{n_d^{-i\sigma}}\rangle\,|\psi_0(\mathbf{n_d^{-i\sigma}})\rangle \\
= \quad & (-1)^{\pi(i\sigma;\mathbf{n_d^{+j\sigma'}})}n_{d,i,\sigma}(-1)^{\pi(j\sigma';\mathbf{n_d})}(1-n_{d,j,\sigma'})\,|\mathbf{n_d^{-i\sigma,+j\sigma'}}\rangle\,|\psi_0(\mathbf{n_d^{-i\sigma,+j\sigma'}})\rangle \\
& + (-1)^{\pi(j\sigma';\mathbf{n_d^{-i\sigma}})}(1-n_{d,j,\sigma'})(-1)^{\pi(i\sigma;\mathbf{n_d})}n_{d,i,\sigma}\,|\mathbf{n_d^{-i\sigma,+j\sigma'}}\rangle\,|\psi_0(\mathbf{n_d^{-i\sigma,+j\sigma'}})\rangle
\end{aligned} \tag{D.16}$$

Again, both summands are equal except for the Fermi sign and cancel each other. Now let $i\sigma = j\sigma'$. We can start the calculation from (*) and arrive at

$$\begin{aligned}
& c_{i\sigma}\Psi_-(i\sigma)(-1)^{\pi(i\sigma;\mathbf{n_d})}(1-n_{d,j,\sigma'})\,|\mathbf{n_d^{+i\sigma}}\rangle\,|\psi_0(\mathbf{n_d^{+i\sigma}})\rangle \\
+ & c^\dagger_{i\sigma}\Psi_+(i\sigma)(-1)^{\pi(i\sigma;\mathbf{n_d})}n_{d,i,\sigma}\,|\mathbf{n_d^{-i\sigma}}\rangle\,|\psi_0(\mathbf{n_d^{-i\sigma}})\rangle
\end{aligned} \tag{D.17}$$

In the next step, we don't have to include the factors $n_{d,i,\sigma}$ and $(1-n_{d,i,\sigma})$ because in this case, they are both equal to 1.

$$\begin{aligned}
= \quad & (-1)^{\pi(i\sigma;\mathbf{n_d^{+i\sigma}})}(-1)^{\pi(i\sigma;\mathbf{n_d})}(1-n_{d,i,\sigma}) \quad |\mathbf{n_d}\rangle\,|\psi_0(\mathbf{n_d})\rangle \\
& + (-1)^{\pi(i\sigma;\mathbf{n_d^{-i\sigma}})}(-1)^{\pi(i\sigma;\mathbf{n_d})}n_{d,i,\sigma} \qquad\quad |\mathbf{n_d}\rangle\,|\psi_0(\mathbf{n_d})\rangle
\end{aligned} \tag{D.18}$$

This time, the Fermi signs are equal. The final result therefore is

$$|\mathbf{n_d}\rangle\,|\psi_0(\mathbf{n_d})\rangle\,. \tag{D.19}$$

Therefore, we have

$$\{\tilde{c}_{i\sigma}, \tilde{c}^\dagger_{j\sigma'}\} = \delta_{i\sigma,j\sigma'}. \tag{D.20}$$

We conclude that the modified particle operators satisfy the usual Fermi anti-commutation relations.

Bibliography

[1] P. A. M. Dirac. Quantum Mechanics of Many-Electron Systems. *Proceedings of the Royal Society of London. Series A.*, 123:714–733, 1929. 1

[2] P.S. Laplace. A Philosophical Essay on Probabilities. Original work published in French in 1814, 1951. 2

[3] P.W. Anderson. More is different. *Science*, 177(4047):393–396, 1972. 2

[4] P.M. Binder. Computation: The edge of reductionism. *Nature*, 459(7245):332–334, 2009. 2

[5] P. Hohenberg, W. Kohn, et al. Inhomogeneous electron gas. *Phys. Rev*, 136(3B):B864–B871, 1964. 3

[6] VV Kohn and LJ Sham. Self-Consistent Equations Including Exchange and Correlation Effects. *PHYSICAL REVIEW*, 140(4A):1–5, 1965. 3, 6

[7] J. Hubbard. Electron Correlations in Narrow Energy Bands. *Proceedings of the Royal Society of London. Series A.*, 276(1365):238–257, 1963. 4

[8] J. Hubbard. Electron correlations in narrow energy bands. ii. the degenerate band case. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 277(1369):237–259, 1964. 4

[9] J. Hubbard. Electron Correlations in Narrow Energy Bands. III. An Improved Solution. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 281(1386):401, 1964. 4

[10] J. Kanamori. Electron correlation and ferromagnetism of transition metals. *Prog. Theor. Phys*, 30(3):275–289, 1963. 4

[11] Martin C. Gutzwiller. Effect of correlation on the ferromagnetism of transition metals. *Phys. Rev. Lett.*, 10(5):159–162, Mar 1963. 4, 50

[12] E.H. Lieb and FY Wu. Absence of Mott transition in an exact solution of the short-range, one-band model in one dimension. *Physical Review Letters*, 20(25):1445–1448, 1968. 4, 42

[13] Walter Metzner and Dieter Vollhardt. Correlated lattice fermions in $d = \infty$ dimensions. *Phys. Rev. Lett.*, 62(3):324–327, Jan 1989. 4

[14] Antoine Georges, Gabriel Kotliar, Werner Krauth, and Marcelo J. Rozenberg. Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. *Rev. Mod. Phys.*, 68(1):13, Jan 1996. 4

[15] M. Born and R. Oppenheimer. Zur quantentheorie der molekeln. *Annalen der Physik*, 389(20):457–484, 1927. 4

[16] N.W. Ashcroft and N.D. Mermin. Solid State Physics, 1976. 4

[17] G.H. Wannier. The structure of electronic excitation levels in insulating crystals. *Physical Review*, 52(3):191–197, 1937. 5

[18] N. Marzari and D. Vanderbilt. Maximally localized generalized Wannier functions for composite energy bands. *Physical review B*, 56(20):12847–12865, 1997. 5

[19] E. Koch and S. Goedecker. Locality properties and Wannier functions for interacting systems. *Solid State Communications*, 119(2):105–109, 2001. 5

[20] F. Lechermann, A. Georges, A. Poteryaev, S. Biermann, M. Posternak, A. Yamasaki, and OK Andersen. Dynamical mean-field theory using Wannier functions: A flexible route to electronic structure calculations of strongly correlated materials. *Physical Review B*, 74(12):125120, 2006. 5

[21] A. Auerbach. *Interacting electrons and quantum magnetism*. Springer, 1994. 8

[22] VI Anisimov, F Aryasetiawan, and AI Lichtenstein. First-principles calculations of the electronic structure and spectra of strongly correlated systems: The LDA+U method. *JOURNAL OF PHYSICS-CONDENSED MATTER*, 9(4):767–808, 1997. 9

[23] F. Aryasetiawan, K. Karlsson, O. Jepsen, and U. Schönberger. Calculations of hubbard u from first-principles. *Phys. Rev. B*, 74(12):125106, Sep 2006. 9, 34, 79

[24] H. Jiang, R.I. Gomez-Abal, P. Rinke, and M. Scheffler. First-principles modeling of localized d states with the GW@ LDA+ U approach. *Physical Review B*, 82(4):45108, 2010. 9

[25] A. I. Lichtenstein, M. I. Katsnelson, and G. Kotliar. Finite-temperature magnetism of transition metals: An ab initio dynamical mean-field theory. *Phys. Rev. Lett.*, 87(6):067205, Jul 2001. 9

[26] William H. Press. *Numerical Recipes: the art of scientific computing*. Cambridge University Press, 3, illustrated edition, 2007. 11, 13, 25

[27] C. Lanczos. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *J. Res. Nat. Bur. Standards*, 45(4):255–282, 1950. 11

[28] A. N. Krylov. On the numerical solution of the equation by which, in technical questions, frequencies of small oscillations of material systems are determined. *Ivz. Akad. Nauk SSSR, ser. fis. mat.*, VII:491–539, 1931. 12

[29] Youcef Saad. *Numerical methods for large eigenvalue problems*. Manchester United Press ND, Manchester, UK, 1992. 12

[30] G.H. Golub and C.F. van Loan. *Matrix Computations*. John Hopkins University Press, Baltimore and London, 3 edition, 1996. 13, 16, 17

[31] Andreas Dolfen. Massively parallel exact diagonalization of strongly correlated systems. Diploma thesis, RWTH Aachen, 2006. 13, 14, 27, 28, 30

[32] S. Kaniel. Estimates for some computational techniques in linear algebra. *Mathematics of Computation*, 20(95):369–378, 1966. 14

[33] C.C. Paige. *The computation of eigenvalues and eigenvectors of very large sparse matrices*. University of London, 1971. 14, 16

[34] CC Paige. Computational variants of the Lanczos method for the eigenproblem. *IMA Journal of Applied Mathematics*, 10(3):373, 1972. 14

[35] CC Paige. Error analysis of the lanczos algorithm for tridiagonalizing a symmetric matrix. *IMA Journal of Applied Mathematics*, 18(3):341, 1976. 16

[36] J.K. Cullum and R.A. Willoughby. *Lanczos Algorithms for Large Symmetric Eigenvalue Computations: Theory*, volume 1. SIAM, 2002. 16

[37] J. Cullum and WE Donath. A block Lanczos algorithm for computing the q algebraically largest eigenvalues and a corresponding eigenspace of large, sparse, real symmetric matrices. In *1974 IEEE Conference on Decision and Control including the 13th Symposium on Adaptive Processes*, volume 13, 1974. 17

[38] G.H. Golub and R. Underwood. The block Lanczos method for computing eigenvalues. *Mathematical software*, 3:361–377, 1977. 17

[39] R.R. Underwood. An iterative block lanczos method for the solution of large sparse symmetric eigenproblems. Technical report, Department of Computer Science, Stanford University, CA, 1975. 17

[40] DS Scott. How to make the lanczos algorithm converge slowly. *Mathematics of Computation*, 33(145):239–247, 1979. 17

[41] E. Dagotto. Correlated electrons in high-temperature superconductors. *Reviews of Modern Physics*, 66(3):763–840, 1994. 18

[42] A.L. Fetter and J.D. Walecka. *Quantum theory of many-particle systems*. Dover Pubns, 2003. 20

[43] R.D. Mattuck. *A guide to Feynman diagrams in the many-body problem*. Dover Pubns, 1992. 20

[44] A. Gupta, A. Grama, G. Karypis, and V. Kumar. *An introduction to parallel computing: design and analysis of algorithms*. Addison Wesley, Reading, MA, 2003. 25, 31

[45] Andreas Dolfen, Yuan Lung Luo, and Erik Koch. Simulating Materials with Strong Correlations on BlueGene/L. In *PARALLEL COMPUTING: ARCHITECTURES, ALGORITHMS AND APPLICATIONS*, volume 15 of *Advances in Parallel Computing*, pages 601–608, 2008. 28

[46] H.P. Breuer and F. Petruccione. *The theory of open quantum systems*. Oxford University Press, USA, 2002. 33

[47] NF Mott and R. Peierls. Discussion of the paper by de Boer and Verwey. *Proceedings of the Physical Society*, 49:72, 1937. 42

[48] Dieter Vollhardt. Normal *he*3: an almost localized fermi liquid. *Rev. Mod. Phys.*, 56(1):99, Jan 1984. 50

[49] Martin C. Gutzwiller. Effect of correlation on the ferromagnetism of transition metals. *Phys. Rev.*, 134(4A):A923–A941, May 1964. 50

[50] F. Aryasetiawan, M. Imada, A. Georges, G. Kotliar, S. Biermann, and A. I. Lichtenstein. Frequency-dependent local interactions and low-energy effective models from electronic structure calculations. *Phys. Rev. B*, 70(19):195104, Nov 2004. 79, 82

[51] M. Springer and F. Aryasetiawan. Frequency-dependent screened interaction in ni within the random-phase approximation. *Phys. Rev. B*, 57(8):4364–4368, Feb 1998. 79

Eigenständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und nur die angegebenen Quellen und Hilfsmittel benutzt habe.

Aachen, den 01. Oktober 2010